HAL
open science

L'évaluation, un objet politique : le cas d'étude de l'aide au développement

Camille Laporte

To cite this version:

Camille Laporte. L'évaluation, un objet politique : le cas d'étude de l'aide au développement. Science politique. Institut d'études politiques de Paris - Sciences Po, 2015. Français. NNT : 2015IEPP0013 . tel-01164807

HAL Id: tel-01164807

https://theses.hal.science/tel-01164807

Submitted on 17 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Institut d'études politiques de Paris

ECOLE DOCTORALE DE SCIENCES PO

Programme doctoral Science politique mention Relations internationales

CERI - Centre d'études de recherches internationales

Doctorat en Science politique

L'évaluation, un objet politique : le cas d'étude de l'aide au développement

Camille Laporte

Thèse dirigée par Guillaume Devin, Professeur des universités à l'IEP de Paris

Soutenue le 23 mars 2015

Jury :

M. Guillaume Devin, Professeur des universités en science politique, IEP de Paris (directeur)

Mme Rosalind Eyben, Emeritus Fellow, Institute of Development Studies, University of Sussex (rapporteur)

M. Jean-Jacques Gabas, Maître de conférences HDR, Université Paris XI (rapporteur)

M. Michel Hastings, Professeur des universités en science politique, IEP de Lille

M. Christian Hugues, Directeur associé de la société EVALUA

Mme Christine Musselin, Directrice de recherche, CNRS-CSO

Remerciements

À l'achèvement de ces trois années et demie de recherche, mes premières pensées vont à mon directeur de thèse, M. Guillaume Devin, que je tiens à remercier pour son soutien, ses encouragements et pour avoir su me donner confiance en mes capacités à faire de la recherche. Je voudrais également le remercier pour son écoute, sa persévérance et sa disponibilité qui m'ont permis de réaliser ce travail le plus sereinement possible.

Ma reconnaissance va également à l'École doctorale de Sciences po Paris pour son soutien aussi bien matériel qu'humain.

Je voudrais ensuite remercier toutes les personnes qui, à travers leurs conseils et les discussions que nous avons pu avoir, m'ont aidée à construire ma réflexion et à améliorer la qualité de cette thèse : Jean-Jacques Gabas, Yves Ficatier, Philippe Hayez, Florent Bédecarrats, Joëlline Bénéfice, Rosalind Eyben et Eric Monnier, pour ne nommer qu'eux. Mes remerciements vont aussi aux membres de mon jury de thèse qui ont accepté de lire et d'évaluer ce travail.

Cette recherche n'aurait pas été aussi enrichissante et épanouissante sans l'appui de l'Agence française de Développement, au sein de laquelle j'ai mené une longue observation participante, de nombreux entretiens et une mission de terrain au Laos. Je tiens, plus particulièrement, à exprimer ma gratitude aux équipes du département de la recherche et des savoirs qui ont été des soutiens précieux de par leur convivialité et leurs partages d'expériences. Ma gratitude va également aux évaluateurs du DFID, de la Banque mondiale, du CAD-OCDE et de la Société française d'évaluation qui m'ont beaucoup appris et qui ont largement contribué à la réussite de cette recherche.

C'est avec générosité et une attention bienveillante que Raphaëlle, Philippe H, Asmara, Isaline, Guillaume et Philippe R se sont attelés à la relecture de ce manuscrit et je leur suis infiniment reconnaissante du temps qu'ils m'ont consacrée.

Si le doctorat est souvent considéré comme un travail solitaire, ce ne fut pas le cas pour moi grâce à la camaraderie de mes amis doctorants, de Sciences po et d'ailleurs, avec qui j'ai pu échanger, nourrir ma pensée et construire des projets de recherche stimulants.

Pour m'accompagner, j'ai aussi bénéficié l'amitié de personnes extérieures au monde de la recherche, dont le soutien a été fondamental dans l'accomplissement de ce travail. Merci à tous mes amis, qui ont été à mes côtés pendant ces trois dernières années pour m'écouter parler des joies et des peines du doctorat et pour les instants de détente qu'ils m'ont fait partager. Merci à ma famille et plus particulièrement à ma sœur, Bérénice, qui a toujours été là pour me soutenir dans mes choix.

Enfin, les dernières années de cette recherche ont été portées par le soutien sans faille, la patience et l'optimisme indéfectibles de Guillaume qui a su m'entourer de tous les soins pour que je puisse conduire ce travail à son terme. Je le remercie infiniment d'avoir été à mes côtés au quotidien et de s'être autant investi dans mes projets.

Table des matières

Acronymes 7
Introduction 11
PARTIE I : Une histoire politique de l'évaluation de l'aide au développement 37
Chapitre 1 : 1960-1991, l'évaluation, un outil de rationalisation de l'aide au développement 41
A. 1960-1978 : l'évaluation, un outil marginal 41
B. Le tournant des années 1980 : l'idéologie néolibérale de l'évaluation 63
C. Les prémices de la normalisation et de l'institutionnalisation de l'évaluation 80
Chapitre 2 : 1991-2000, l'évaluation, un outil de légitimation de l'aide au développement 95
A. L'évaluation comme élément de réponse à la perte de légitimité de l'aide 97
B. Élargissement et complexification de la fonction d'évaluation 105
C. Des standards de l'évaluation influencés par la doctrine du New Public Management 117
Chapitre 3 : 2000-2011, l'évaluation, un outil de dépolitisation de l'aide au développement 145
A. Des approches contradictoires de l'évaluation 146
B. Une demande politique croissante d'évaluations expérimentales 159
C. Des « bonnes pratiques » de l'évaluation difficiles à opérationnaliser 171
PARTIE II : La diffusion d'une culture de l'évaluation : acteurs, leviers et réseaux 187
Chapitre 4 : Les évaluateurs, des acteurs peu reconnus 191
A. Une négation des dynamiques interactionnelles 193
B. Jeux d'acteurs et rapports de pouvoir 212
C. Utilité et usages de l'évaluation 231
Chapitre 5 : L'artifice du renforcement des capacités en évaluation au Sud 241
A. Des transferts de compétences inopérants 241
B. Les évaluations partenariales, un vecteur d'apprentissage mal exploité 263
C. Des systèmes d'évaluation désincarnés 279
Chapitre 6 : Les réseaux d'évaluateurs : quelle influence normative et politique dans le régime de l'aide ? 293
A. Des réseaux instrumentalisés par les organisations multilatérales 293
B. Donner une existence politique aux évaluateurs 310
C. Une faible influence sur les gouvernements du Sud 325
PARTIE III : Les effets de l'évaluation : reclassement, contrôle ou démocratisation ? 337
Chapitre 7 : Un régime de l'aide fragmenté autour des « bonnes pratiques » d'évaluation : le catalyseur des donateurs émergents 343

A. Des effets de déclassement sur la scène internationale 343
B. Une influence normative réciproque 355
C. Des divisions politiques au niveau local 375
Chapitre 8 : Les effets de l'évaluation au Sud, entre contrôle et démocratisation 391
A. Une allocation de l'aide davantage fondée sur la bonne gouvernance que sur la performance 391
B. L'évaluation, entre contrôle et démocratisation des gouvernements et des populations 403
C. Remise en question de l'existence d'une « redevabilité mutuelle » entre donateurs et bénéficiaires de l'aide 424
Chapitre 9 : Les effets de reclassement de l'évaluation des donateurs 441
A. Des politiques d'aide des donateurs bilatéraux encore opaques 442
B. Une évaluation grandissante de l'efficacité des organisations multilatérales 452
C. Un instrument technique aux effets politiques 471
Conclusion 485
Annexe 1 : Cadre logique 497
Annexe 2 : Approches de l'aide et conditionnalités sous-jacentes 498
Annexe 3 : Revue des principales études académiques sur l'efficacité de l'aide (1970-2012) 499
Annexe 4 : Émergence des réseaux d'évaluateurs 501
Annexe 5 : Excuses pour éviter l'évaluation 502
Annexe 6 : Logiques d'allocation de l'aide des donateurs 503
Annexe 7 : Lexique 504
Annexe 8 : Liste des entretiens menés 509
Annexe 9 : Liste non exhaustive des évènements observés 513
Annexe 10 : Profils d'évaluateurs 517
Bibliographie 524

Acronymes

ACDI : Agence canadienne de développement international

AFD : Agence française de développement

AFREA : Association africaine d'évaluation

APD : Aide publique au développement

AUSAID : Agence d'aide australienne

BIRD : Banque internationale pour la reconstruction et le développement

CAD-OCDE : Comité d'aide au développement de l'Organisation de coopération et de développement économique

CE : Commission européenne

CGDev : Center for global development

CICR : Comité internationale de la Croix Rouge

CLEAR : Centre régional pour l'apprentissage des résultats d'évaluations

COMPAS : Common performance assistance system

CPIA : Country policy and institutional assessment

CPR : Country performance rating

DANIDA : Agence d'aide au développement danoise

DFID : Department for international development

ECG : Evaluation cooperation group

ECHO: Office d'aide humanitaire de la Commission européenne.

EDCF : Fonds de coopération pour le développement économique

FAO : Organisation des Nations unies pour l'alimentation et l'agriculture

FIDA : Fonds international de développement agricole

FMI : Fonds monétaire international

FMSTP : Fonds mondial de lutte contre le sida, le paludisme et la tuberculose

GAVI : Alliance globale pour la vaccination et l'immunisation

GCDT : Fonds global pour la diversité des cultures agricoles

GCIAR : Groupe consultatif pour la recherché agricole internationale

GFDRR : Global facility for disaster reduction and recovery

IATI : International aid transparency initiative

IDA : Association pour le développement international (Banque mondiale)

IDEAS : Association pour le développement international de l'évaluation

IDH : Indice de développement humain

IEG : Groupe d'évaluation indépendant (Banque mondiale)

IFRC: Fédérations internationales de la Croix Rouge et du Croissant Rouge

IOCE : Organisation internationale pour la coopération dans l'évaluation

IOM : Organisation internationale pour les migrations

IPDET : International program for development evaluation training

IPEN : International program evaluation network

IPRCC : International poverty reduction center in China

J-PAL : The Abdul Latif Jameel poverty action lab

JICA : Agence japonaise de coopération pour le développement

KOICA : Agence coréenne de coopération pour le développement

LDCF : Fonds pour les pays les moins avancés

MEPP : Mission d'étude, d'évaluation et de prospective

MOPAN : Multilateral organisation performance assessment network

NEI : Need-effectiveness index

OCHA : Office des Nations unies pour la coordination de l'action humanitaire

ODA : Organisation development administration

ODI : Oversea development institute

OED : Operation evaluations department

OIT: Organisation international du travail

OMD : Objectifs du millénaire pour le développement

OMS : Organisation mondiale de la santé

ONG : Organisation non gouvernementale

ONUDI: Organisation des Nations unies pour le développement industriel

OSC : Organisations de la société civile

PAM : Programme alimentaire mondial

PBF : Fonds des Nations unies pour le maintien de la paix

PIB : Produit intérieur brut

PIDG : Private infrastructure development group

PNB : Produit national brut

PNUD : Programme des Nations unies pour le développement

PPBS : Program, planning and budgeting system

QUODA : Quality of development aid assessment

RCB : Rationalisation des choix budgétaires

RGPP : Révision générale des politiques publiques

RRA : Rural rapid assessment

SIDA : Swedish international development agency

SFE : Société française d'évaluation

UNCTAD : United Nations conference on trade and development

UNHCR : Haut-Commissariat des Nations unies pour les réfugiés

UNICEF : Fonds des Nations unies pour l'enfance

UNISDR : Office des Nations unies pour la réduction des risques naturels

UNEP : Programme des Nations unies pour l'environnement

UNODC : Office des Nations unies contre la drogue et la corruption

UNRWA : Office de secours et de travaux des Nations unies pour les réfugiés de Palestine dans le Proche-Orient

USAID : United States agency for international development

3IE : Initiative internationale pour les évaluations d'impact

Introduction

L'année 2015 a été déclarée « année internationale de l'évaluation » par le réseau d'évaluateurs Evalpartners, un partenariat international créé en 2010 pour promouvoir la culture de l'évaluation à travers le monde, et plus spécifiquement dans les pays du Sud. Evalpartners est une plateforme de 34 organisations incluant de nombreuses associations professionnelles d'évaluateurs, des organisations multilatérales, le réseau d'évaluateurs du Comité d'aide au développement de l'organisation de coopération et de développement économique (CAD-OCDE), des organismes d'aide bilatéraux, des fondations privées et des ONG. L'initiative Evalyear est portée par Marco Segone, un évaluateur ayant travaillé dans plusieurs organisations spécialisées des Nations unies, qui s'investit depuis des décennies dans le renforcement des capacités d'évaluation des pays bénéficiaires de l'aide au développement. Il est assisté par Natalia Kosheleva et Jim Rugh, deux évaluateurs expérimentés et particulièrement impliqués dans les réseaux transnationaux d'évaluateurs qui se multiplient depuis le début des années 2000.

Dans des entretiens menés en 2014[1], ces acteurs expliquent que l'initiative Evalyear ambitionne d'être le vecteur d'un plaidoyer en faveur de l'évaluation destiné aux acteurs internationaux, régionaux, nationaux et locaux œuvrant pour un développement durable et équitable. Cette initiative doit permettre de diffuser l'idée que l'évaluation peut être un outil d'appui à la bonne gouvernance pour les décideurs, ainsi que d'*empowerment*[2] pour la société civile. Les porteurs de cette initiative veulent montrer que l'évaluation est un instrument de démocratisation des politiques de développement, c'est-à-dire qu'elle permet de rendre ces politiques plus transparentes, plus ouvertes, plus efficaces et plus pertinentes par rapport aux besoins des populations. Enfin, ils soulignent leur ambition de mettre la réflexion sur l'évaluation au centre des enjeux de la coopération pour le développement après 2015.

La date de 2015 n'a pas été choisie au hasard. Cette année marquera la fin des Objectifs du millénaire pour le développement (OMD) qui seront remplacés par de nouveaux objectifs

[1] Entretien avec la coresponsable du réseau d'évaluateurs IOCE, 22 janvier 2014 ; entretien avec un évaluateur au sein des Nations unies, 6 octobre 2014.

[2] L'empowerment est un concept très utilisé dans le secteur de l'aide au développement, mais qui reste mal définie. Dans le cadre de notre recherche sur l'évaluation, nous comprenons ce terme comme une finalité des programmes de renforcement des capacités en évaluation de la société civile afin qu'elle gagne en influence politique pour inciter les gouvernements à engager des réformes visant à davantage de transparence et de démocratie participative. Pour une définition plus générale de ce terme voir lexique en annexe 7 ou se référer à : CALVES, Anne-Emmanuèle, « Empowerment : généalogie d'un concept clé du discours contemporain sur le développement », *Revue Tiers monde,* vol.4, n°200, 2009, pp. 735-749.

internationaux : les objectifs de développement durable. L'enjeu pour les évaluateurs est de s'assurer que l'évaluation soit considérée comme un instrument central permettant de progresser vers l'atteinte des objectifs de développement durable et des objectifs nationaux de développement qui les accompagneront.

Établir une « année internationale de l'évaluation » permet de poser un cadre à l'organisation de multiples évènements et rencontres autour de ce sujet, de donner de la visibilité à l'évaluation, de créer une communauté d'évaluateurs et de diffuser une pensée et un discours communs sur l'évaluation. Cela permet également de montrer que l'évaluation est un enjeu central des politiques de développement, à tel point qu'elle mérite de se voir consacrer une année de débats et de réflexion. Alors que l'évaluation n'était qu'une pratique marginale de l'aide au développement dans les années 1970, elle mobilise aujourd'hui un grand nombre d'acteurs du développement qui s'en saisissent et tentent de se l'approprier.

Il est intéressant de se pencher sur l'image de l'évaluation qu'essaient de véhiculer les organisateurs de l'initiative « Evalyear ». Ils la présentent comme un instrument au service de la société civile et des gouvernements des pays du Sud permettant de rendre plus équitable et ouverte la mise en œuvre de la coopération internationale pour le développement. Or, les pratiques d'évaluation dans le cadre de cette coopération ont été institutionnalisées sous l'égide du CAD-OCDE dans les années 1990, marquées par une crise de l'aide due au constat que la décennie précédente avait été « perdue » pour l'aide au développement. L'évaluation répondait à une nécessité de « rationalisation » de cette aide. Elle a été conçue comme un outil permettant d'améliorer la redevabilité des politiques d'aide, de rationaliser l'allocation des fonds et de gérer les programmes d'aide selon une logique de résultats et non plus de moyens. C'est ainsi qu'évaluation, redevabilité et gestion axée sur les résultats ont composé un triptyque au service de l'efficacité de l'aide.

Afin que l'efficacité de l'aide soit évaluée de manière harmonisée dans tous les contextes et par l'ensemble des acteurs du développement, l'évaluation a été l'objet d'un processus de normalisation. Toute évaluation doit, en effet, répondre à un certain nombre de standards, de règles et de codes qui en garantissent la qualité. L'évaluation est ainsi devenue une bonne pratique de l'aide au développement, c'est-à-dire un passage obligé de tout programme d'aide et une pratique systématisée et institutionnalisée au sein des organismes d'aide.

Cette conception bureaucratique de l'évaluation va à l'encontre de l'image diffusée par les initiateurs d'« Evalyear » qui présentent l'évaluation comme un outil de gouvernement souple, pouvant s'adapter aux différents contextes et capacités des pays du Sud. Elle apparaît,

en effet, difficilement conciliable avec une conception de l'évaluation qui en ferait un instrument au service des pays bénéficiaires de l'aide pour qu'ils s'emparent des programmes d'aide au développement et deviennent autonomes dans la gestion de leurs politiques publiques. L'évaluation étant, aujourd'hui, menée pour et par les donateurs, les pays bénéficiaires de l'aide ont toujours tendance à considérer qu'elle est un outil qui leur est imposé et qui bureaucratise les programmes d'aide, afin de servir les besoins de redevabilité des donateurs. C'est ce qui conduit les défenseurs d'une approche participative et émancipatrice de l'évaluation à mener une entreprise internationale de plaidoyer pour démontrer que l'évaluation peut être un instrument au service des gouvernements et des sociétés civiles du Sud.

Cette tension entre une approche « technocratique » de l'évaluation, au service des besoins de redevabilité des organismes d'aide, et une conception « démocratique » de l'évaluation, au service de l'appropriation des programmes d'aide au développement par leurs bénéficiaires, a toujours été latente dans le régime de l'aide au développement. Elle monte en puissance aujourd'hui, dans un contexte où les équilibres entre donateurs traditionnels, donateurs émergents, donateurs non-gouvernementaux (ONG, fondations privées) et bénéficiaires de l'aide se redéfinissent.

L'évaluation ne se réduit pas un instrument neutre de gestion de l'aide au développement. Les fonctions de l'évaluation relèvent donc en grande partie du politique, c'est-à-dire de la régulation[3] de la collaboration entre des individus aux intérêts multiples pour les guider vers le bien commun : des politiques d'aide qui permettent un développement durable et équitable au Sud. L'évaluation est mobilisée par les acteurs de l'aide pour prouver qu'ils produisent des résultats et donc légitimer leurs actions, pour améliorer leurs politiques et stratégies, pour contrôler la bonne mise en œuvre des programmes d'aide, ou encore pour intégrer les différentes parties prenantes des programmes d'aide dans un processus d'apprentissage et de compréhension mutuel. Les enjeux de l'évaluation sont déterminés par les débats plus généraux sur l'efficacité de l'aide, ainsi que par les multiples rapports de force qui se construisent dans le cadre de l'aide au développement, que ce soit entre donateurs et bénéficiaires, entre donateurs traditionnels et donateurs émergents, ou entre les donateurs traditionnels eux-mêmes. C'est cette approche de l'évaluation, en tant qu'objet politique de l'aide au développement, qui sera développée dans cette recherche.

[3] Nous entendons par régulation, les processus de contrôle par lesquels un ensemble ou un mécanisme social donné maintient un minimum d'ordre et d'intégration. HERMET *et al.*, Dictionnaire de la science politique, Armand Colin, 6ème éd. 2005.

- **Champ de la recherche et définitions**

L'évaluation des politiques publiques est une pratique qui a émergé aux États-Unis dans les années 1960 et s'est développée dans les années 1980 dans le régime de l'aide au développement. Selon Michael Scriven et Jane Roth, en 1991, « l'évaluation est le processus par lequel sont déterminés le mérite, l'importance et la valeur des programmes de politiques publiques »[4]. Carol Weiss, en 1998, définit l'évaluation comme « l'analyse de la mise en œuvre et/ou la mesure des résultats d'un programme ou d'une politique publique en fonction d'un ensemble de standards implicites ou explicites, afin de contribuer à l'amélioration de ce programme ou de cette politique »[5]. En 2003, la société française d'évaluation propose une définition de l'évaluation en ces termes : « l'évaluation est une activité d'étude et d'analyse portant sur la mise en œuvre et les résultats d'une action, menée dans un cadre méthodologique et institutionnel formalisé dans le but de former des jugements empiriquement et normativement fondés sur la pertinence, la cohérence, l'efficacité et/ou l'efficience de cette action »[6]. En 2008, Michael Patton décrit, quant à lui, l'évaluation comme « la collecte systématique d'informations sur les activités, les caractéristiques et les résultats de programmes de politiques publiques, afin d'établir des jugements sur ces programmes, d'améliorer leur efficacité et/ou d'informer la prise de décision sur de futures programmations »[7].

De ces définitions ressortent plusieurs points communs : l'évaluation est perçue comme un instrument d'analyse des résultats d'un programme de politique publique qui permet de fournir un jugement sur l'efficacité de ce programme et de l'améliorer. Toutefois, il n'a jamais existé une définition unique et stable de l'évaluation. Ainsi, alors que certains considèrent qu'il revient aux évaluateurs de porter un jugement sur les programmes de politiques publiques évalués, d'autres estiment que cette tâche revient aux destinataires de l'évaluation à partir des constats établis. Certains évaluateurs estiment aussi que le but premier d'une évaluation est la mesure de l'efficacité d'une politique publique, tandis que d'autres favorisent la fonction d'apprentissage que peut permettre une analyse qualitative de la mise en œuvre de cette politique. Beaucoup mêlent ces différents aspects.

---

[4] SCRIVEN, Michael, ROTH, Jane, *Evaluation thesaurus*, Edgepress, 1991, 391p. pp. 11-12.
[5] WEISS, Carole, "Have we learned anything new about the use of evaluation?", *American Journal of Evaluation*, vol. 19, n°1, 1998.
[6] Société française d'évaluation, *Charte de l'évaluation des politiques publiques et des programmes publiques,* 28 octobre 2003.
[7] PATTON, Michael, Q, *Utilization focused evaluations*, Sage Publication, fourth Edition, 2008, 688p. p. 39.

Le champ de l'évaluation varie donc selon les approches des évaluateurs et la demande des décideurs politiques. L'évaluation peut se restreindre à l'analyse des réalisations ou des activités d'un programme ou d'une politique, s'étendre à l'analyse de leurs résultats (donc de leurs effets directs, de court terme et sur un public ciblé) ou encore aller jusqu'à étudier leurs impacts (c'est-à-dire leurs effets de long terme et indirects sur une population plus large que les bénéficiaires ciblés). L'évaluation est donc un concept souple et à géométrie variable, ce qui rend difficile son appréhension.

Cerner l'évaluation nécessite également de la distinguer des disciplines qui lui sont proches, à savoir l'audit, le contrôle interne et le suivi. En dehors du cercle des évaluateurs, les acteurs de l'aide au développement interrogés dans le cadre de notre recherche témoignent d'une difficulté encore réelle à définir les particularités de l'évaluation par rapport à ces activités. L'audit consiste à contrôler la légalité, la conformité et la régularité de l'utilisation de ses ressources par une organisation. Il est réalisé de manière périodique et selon des normes préétablies dans les statuts légaux de l'organisation. L'audit sert à déceler et prévenir des irrégularités dans la gestion organisationnelle et financière d'une organisation. Il peut être réalisé par des auditeurs externes ou internes, ou encore par la Cour des comptes[8]. Contrairement à l'évaluation, l'audit produit des recommandations qui ont valeur obligatoire devant la loi. L'évaluation, quant à elle, fait des recommandations qui ont valeur de conseil. Si elle peut s'intéresser à la gestion des programmes de politiques publiques et à l'utilisation des ressources financières qu'ils mobilisent, elle s'inscrit dans un champ plus large qui s'étend à l'analyse de l'efficacité, de l'efficience, de la pertinence, de l'utilité et de la cohérence de ces programmes.

Par ailleurs, le contrôle de gestion vise à assurer la conformité d'une organisation aux lois et aux règlements, l'application des instructions et des orientations fixées par la direction générale ou le directoire, le bon fonctionnement des processus internes de l'organisation et la fiabilité des informations financières. Le contrôle se réfère à des normes internes au système analysé (règles comptables, juridiques, ou normes fonctionnelles), tandis que l'évaluation essaie d'appréhender, d'un point de vue principalement externe, les effets et/ou la valeur de l'action considérée[9].

---

[8] Nous avons déterminé ces distinctions à partir des définitions de l'audit, du contrôle et de l'évaluation données par les organismes suivants : l'Institut français de l'audit et du contrôle interne : http://www.ifaci.com/ifaci/connaitre-l-audit-et-le-controle-interne/definitions-de-l-audit-et-du-controle-internes-78.html, consulté le 18 décembre 2014 ; du Sénat, http://www.senat.fr/rap/r03-392/r03-3922.html, consulté le 18 décembre 2014 ; et de la Cour des comptes, *Rapport public annuel 2012, Chapitre 3 « les première évaluations des politiques publiques »*, 2012.

[9] *Ibid.*

Il existe, cependant, des zones de recoupement. C'est le cas lorsque la norme d'efficacité est prise en compte par le contrôle interne, ou lorsque l'évaluation ne se contente pas de mesurer l'efficacité d'un programme ou d'une politique, mais aussi son efficience, c'est-à-dire le ratio entre les moyens et les finalités de l'action publique. L'évaluation appelle d'autres compétences que le contrôle et, plus encore, la participation d'une gamme beaucoup plus étendue d'intervenants.

Enfin, le suivi est un processus continu de collecte systématique d'informations pour fournir aux parties prenantes d'une politique publique en cours des éléments sur les progrès réalisés, les objectifs atteints et l'utilisation des fonds alloués. Alors que l'évaluation a lieu une fois le programme de politique publique achevé, le suivi se fait de manière continue durant le projet. Le suivi n'est pas un exercice analytique. Il se limite à la collecte de données qualitatives et quantitatives qui vont ensuite servir à l'évaluation. Il est donc possible de considérer le suivi comme l'étape préalable à l'évaluation, conditionnant sa bonne réalisation. Les termes « suivi » et « évaluation » ont tendance à être assimilés dans l'expression « suivi-évaluation » qui présente les deux activités comme un continuum. Les évaluateurs s'opposent souvent à cet amalgame, considérant qu'il réduit l'évaluation à un outil gestionnaire et tend à occulter son aspect d'analyse des politiques publiques et de production de connaissances. Toutefois, cette expression a l'avantage de montrer que l'évaluation est une activité qui doit se fonder sur des données quantitatives et des informations qualitatives solides sur les programmes considérés. Une évaluation ne peut être mise en œuvre de manière rigoureuse que si elle repose sur un travail préalable de suivi exhaustif et continu[10].

Ce travail de définition montre que la nuance entre évaluation, audit, contrôle et suivi est fine. L'évaluation étant une activité plus récente et moins institutionnalisée que l'audit et le contrôle, elle a longtemps peiné à se positionner et à se faire reconnaître en tant qu'expertise à part entière des politiques de développement. Pour pallier ce flou et cette ambiguïté autour de la notion d'évaluation, le groupe d'évaluateurs du CAD-OCDE a produit, en 1991, une définition de l'évaluation de l'aide au développement, afin que tous les acteurs de ce secteur d'activité en aient la même approche : « appréciation systématique et objective d'un projet, d'un programme ou d'une politique, en cours ou terminé, de sa conception, de sa mise en œuvre et de ses résultats. Le but est de déterminer la pertinence et l'accomplissement des objectifs, l'efficience en matière de développement, l'efficacité, l'impact et la durabilité. Une évaluation devrait fournir des informations crédibles et utiles permettant d'intégrer les leçons de

[10] Cette distinction est issue de documents de formation internes à l'Agence française de développement.

l'expérience dans le processus de décision des bénéficiaires et des bailleurs de fonds. Le terme "évaluation" désigne également un processus aussi systématique et objectif que possible par lequel on détermine la valeur et la portée d'une action de développement projetée, en cours ou achevée »[11]. Cette définition a été communément adoptée par les donateurs bilatéraux et multilatéraux et est, aujourd'hui encore, considérée comme la définition officielle de l'évaluation dans le cadre de l'aide au développement.

D'après cette définition, une évaluation peut avoir lieu en cours de programme, ou encore une fois ce dernier achevé. En effet, si l'évaluation est le plus souvent réalisée à la fin d'un projet pour en mesurer l'efficacité, c'est-à-dire l'atteinte des objectifs qui lui étaient fixés, elle peut également prendre place à mi-parcours, répondant alors à une volonté formative, c'est-à-dire d'apprentissage et d'amélioration du projet. Ainsi, si l'évaluation à mi-parcours fait ressortir des difficultés ou des manquements, les opérateurs du projet sont en capacité de les traiter et de réorienter le projet. Un autre type d'évaluation existe également, l'évaluation *ex ante,* conduite avant la mise en œuvre d'un projet ou d'un programme. Elle évalue la faisabilité et les risques d'un projet envisagé par des organismes d'aide. Dans le cadre de cette recherche, nous avons choisi de nous focaliser sur les évaluations *ex post* des programmes d'aide au développement. En effet, les évaluations *ex ante* sont des exercices très différents qui posent des enjeux politiques d'un autre ordre lesquels mériteraient une recherche à part entière[12].

- **Périmètre de la recherche**

L'aide au développement apparaît comme un cadre particulièrement intéressant pour analyser l'objet politique qu'est l'évaluation. Il est tout d'abord frappant de constater qu'alors que l'évaluation était une pratique marginale jusque dans les années 1980, elle est aujourd'hui un passage obligé de toute activité de coopération pour le développement. Les projets et les programmes d'aide sont systématiquement évalués par tous les donateurs membres du CAD-OCDE. Ainsi, l'institutionnalisation de l'évaluation dans les politiques d'aide au développement est un phénomène international généralisé.

Ce processus de normalisation et de systématisation de l'évaluation dans les organismes d'aide au développement a été impulsé par des « opérateurs de normes » au niveau multilatéral qui ont incité les donateurs bilatéraux à pratiquer l'évaluation selon les standards uniformisés

[11] DAC-OECD, *Glossaire du CAD de l'OCDE sur les principaux termes relatifs à l'évaluation*, 1991.
[12] Nous pourrons toutefois faire ponctuellement référence à ce type d'évaluation si cela est pertinent dans le cadre de notre argumentation.

qu'ils ont édictés. L'évaluation de l'aide au développement est ainsi devenue une bonne pratique partagée par l'ensemble des donateurs occidentaux. Par bonne pratique, nous entendons « un ensemble de règles, instructions ou codes de conduite fondés sur l'idée qu'il existe, pour atteindre un résultat donné, une méthode meilleure que les autres. Issues de l'expérience passée au filtre de l'étude statistique, les bonnes pratiques sont donc à la fois un standard qui prétend décrire objectivement les voies de la performance, en même temps qu'un guide pour l'action. En somme, les bonnes pratiques ont pour vocation de normaliser les comportements, afin de réduire l'écart entre le fait et le prescrit »[13]. L'évaluation de l'aide au développement est une bonne pratique, dans la mesure où elle répond à un certain nombre de règles déterminées par le CAD-OCDE et la Banque mondiale, qui se sont imposés comme les organisations productrices de normes en la matière[14]. De plus, elle remplit bien le rôle d'une bonne pratique qui est de guider les acteurs vers les voies de la performance.

La « bonne pratique »[15] de l'évaluation s'est construite à partir de l'expérience des évaluateurs pour les guider dans leur tâche et répondre aux difficultés qu'ils pouvaient rencontrer. Si la pratique de l'évaluation est donc rendue routinière par des standards et des procédures qui guident l'action des évaluateurs, elle continue d'évoluer à mesure que les approches de l'efficacité de l'aide prônées par les organisations multilatérales, telles que les Nations unies, la Banque mondiale, le CAD-OCDE ou la Commission européenne se sont transformées Les bonnes pratiques sont des objets souples qui se transforment en fonction des besoins des acteurs et des organisations qui les emploient, ainsi que du contexte dans lequel elles sont utilisées.

Il faut ensuite remarquer que dans un grand nombre de pays donateurs, la pratique de l'évaluation est plus développée dans le secteur de l'aide au développement que dans les autres secteurs des politiques publiques. S'en suit alors un processus de transfert des « bonnes pratiques » de l'évaluation de l'APD vers d'autres secteurs. C'est le cas, par exemple, en France, en Espagne ou au Japon. Étudier l'évaluation à travers l'aide au développement permet donc de comprendre et d'anticiper les évolutions de la discipline de l'évaluation dans les autres politiques publiques.

[13] KLEIN, Asmara, LAPORTE, Camille, SAIGET, Marie, *Les bonnes pratiques des organisations internationales*, Paris, Presses de Sciences po, à paraître en mars 2015.
[14] Les codes d'éthique élaborés par les sociétés nationales d'évaluation reprennent en très grande majorité les standards du CAD-OCDE.
[15] Nous mettons le terme bonne(s) pratique(s) entre guillemets lorsqu'il détermine un objet particulier, c'est-à-dire qu'il relève d'une construction politique. Lorsqu'il est utilisé comme un concept générique, nous n'employons pas de guillemets.

La particularité de la politique publique de l'aide au développement est, par ailleurs, que ses bénéficiaires (les populations des pays du Sud) se distinguent de ses financeurs (les contribuables de l'État donateur), alors que dans les autres politiques publiques ces deux catégories d'acteurs se confondent. En conséquence, l'évaluation devient l'instrument fondamental qui va permettre aux contribuables et aux décideurs politiques d'un État donateur d'avoir un retour rigoureux et indépendant sur les résultats des politiques d'aide qui s'adressent à un public lointain. L'évaluation permet de faire la connexion entre les décideurs politiques, les contribuables, les opérateurs de l'aide et les bénéficiaires de l'aide. Ces différents acteurs ont des perceptions différentes de l'efficacité de l'aide qui se confrontent dans l'exercice de l'évaluation[16]. Les évaluations révèlent les tensions latentes dans l'aide au développement entre les attentes des Etats bénéficiaires de l'aide, les besoins de leur population, les priorités et stratégies des décideurs politiques des États donateurs et les approches de l'aide et les contraintes organisationnelles des agences d'aide.

- **État de l'art**

Si l'évaluation est devenue un aspect incontournable de l'aide au développement, elle reste un objet peu étudié par la littérature sur ce thème. Les rares ouvrages qui traitent spécifiquement de la question de l'évaluation appliquée au secteur de l'aide au développement sont anciens (plus de dix ans), se restreignent largement à une analyse de l'évaluation dans ses aspects méthodologiques, proviennent des pays anglo-américains et sont le plus souvent rédigés par des évaluateurs et non par des chercheurs en sciences sociales[17]. Il n'existe aucune étude présentant une histoire politique globale et actualisée de l'évaluation de l'aide au développement et offrant une analyse générale des enjeux politiques soulevés par l'évaluation dans le régime de l'aide au développement.

Les ouvrages généraux sur l'aide au développement n'évoquent pas, ou n'abordent que succinctement, l'évaluation[18]. Rares sont ceux qui consacrent plusieurs paragraphes aux enjeux liés à l'évaluation[19]. Face à ce manque de recherches académiques et au vu de la montée en

[16] Selon le nombre de ces acteurs inclus dans le processus évaluatif.
[17] Voir bibliographie, section « évaluation de l'aide au développement ».
[18] RIST, Gilbert, *Le développement, histoire d'une croyance occidentale*, Presses de Sciences po, 4ème édition, 2013, 512p ; SEVERINO, Jean-Michel, DEBRAT, Jean-Michel, *L'aide au développement,* Le Cavalier Bleu, Paris, 2009, 127p ; PHILLIPS, David. A, *Development without Aid, the Decline of Development Aid and the Rise of Diaspora*, Anthern Press, 2013, 223p; LECOMTE, Bernard, NAUDET, Jean-David, *Survivre grâce à l'aide, réussir malgré l'aide*, Cahier des Sciences humaines, n°13, Editions de l 'Aube, 2000, 211p ; Mc NAMARA, Robert, *One Hundred Countries, Two Billions People: the dimensions of Development,*: Praeger, 1973, 140p.
[19] EASTERLY, William, *The White Man's Burden: why the West's efforts to aid the rest have done so much ill and so little good,* The Penguin Press, 2006, 436p; GABAS, Jean-Jacques, *Nord-Sud : l 'impossible cooperation,* Presses de Sciences po,

puissance de l'évaluation dans le secteur de l'aide au développement, il nous a paru pertinent de réaliser une recherche ayant comme objet principal l'évaluation.

En revanche, les travaux académiques entièrement consacrés à la question de savoir si l'aide est efficace ou non sont nombreux, et ce depuis des décennies[20]. Ainsi, si beaucoup de chercheurs se sont attachés à analyser les facteurs de réussite et d'échec de l'aide au développement, peu d'entre eux ont tenté de comprendre les sous-jacents politiques qui déterminent l'évaluation de cette efficacité. Or, l'efficacité de l'aide est loin d'être un concept neutre et objectif faisant l'objet d'un consensus. La preuve en est que les ouvrages portant sur ce thème présentent des considérations variées sur l'efficacité de l'aide et les moyens de l'atteindre[21]. Nous voyons l'efficacité comme une notion évolutive, qui se transforme à mesure que les priorités politiques des donateurs changent, que les relations entre les donateurs et les bénéficiaires se modifient et que le savoir sur le développement s'accumule. L'évaluation est l'un des outils qui permettent de construire une image de ce qu'est une aide efficace. En effet, les standards, les objectifs et les indicateurs sur lesquels se fondent les évaluations tendent à objectiver les critères de l'efficacité de l'aide, telle qu'elle est perçue à un moment donné de l'histoire du développement. Ainsi évaluation et efficacité de l'aide sont deux enjeux étroitement liés qu'il est impossible de traiter indépendamment. Les approches de l'efficacité de l'aide conditionnent la pratique de l'évaluation, qui à son tour conduit à les pérenniser ou à les faire évoluer. Nous observons ainsi un phénomène de rétroaction entre efficacité de l'aide et évaluation.

Le faible intérêt de la littérature académique pour la question de l'évaluation dans le secteur de l'aide au développement contraste avec la profusion de la littérature grise produite à ce sujet par les organismes de développement. Les papiers de recherche, guides méthodologiques, cadres normatifs et stratégies politiques réalisés par l'OCDE[22], la Banque

2002, 114p ; DANN, Philip, *The Law of Development Cooperation, A Comparative Analysis of the World Bank, the EU and Germany*, Cambridge University Press, 2013, 604p ; DEGNBOL-MARTINUSSEN, John, ENGBERG-PEDERSEN, Paul, *Aid: Understanding International Development Cooperation*, Zed Books, 2005, 353p.

[20] Voir annexe 3.

[21] Voir annexe 3.

[22] Rien que sur le site du CAD-OCDE, on compte plus de 150 documents traitant de la thématique de l'évaluation : http://www.oecd.org/fr/cad/evaluation/latestdocuments/ (consulté le 24 septembre 2014) dont certains sont cités en bibliographie.

mondiale[23] et les organisations des Nations unies[24] ou de la Commission européenne[25] se comptent par centaines. Chaque agence d'aide bilatérale membre du CAD-OCDE dispose de son propre document de politique d'évaluation et de ses propres guides méthodologiques ; nombreuses sont celles à produire régulièrement des études sur l'évaluation.

La littérature portant sur les évolutions des approches et des méthodes d'évaluation des politiques publiques est un corpus de référence essentiel pour un travail de recherche sur l'évaluation de l'aide au développement. Il y a tout d'abord les travaux cherchant à améliorer les techniques d'évaluation en fonction des attentes des acteurs des politiques publiques et des avancées en sciences sociales et de la méthodologie de l'évaluation[26]. Ces travaux sont intéressants, mais ils ne permettent pas de comprendre les enjeux politiques sous-jacents aux débats et aux choix méthodologiques sur l'évaluation, enjeux qui sont l'objet d'étude central de cette recherche. D'autres travaux développent une approche culturelle de l'évaluation, replaçant les pratiques d'évaluation dans le contexte politique et culturel dans lequel elles sont mises en place. Il s'agit généralement de travaux de politique comparée analysant les variations de l'institutionnalisation et de la pratique de l'évaluation entre un certain nombre de pays, à l'aune des différences de culture politique qui les caractérisent[27]. Notre recherche s'inspire, en partie, de ce courant, en déconstruisant l'idée selon laquelle l'évaluation serait une activité neutre et harmonisée parmi les acteurs de l'aide au développement. Nous démontrerons que les pratiques d'évaluation divergent, en effet, d'un organisme d'aide à un autre en fonction de ses contraintes organisationnelles et de sa culture politique. La pratique de l'évaluation de l'aide au développement est donc sous la triple influence des organismes multilatéraux normalisateurs, du contexte politique international de l'aide au développement et de la culture des politiques publiques nationales (pour les agences d'aide bilatérales).

---

[23] Sur le site de *l'Independent Evaluation Group* de la Banque mondiale, on trouve environ 250 travaux de réflexion sur l'activité d'évaluation de la Banque : http://ieg.worldbankgroup.org/ieg-search?f%5B0%5D=field_document_categories%3A343, consulté le 24 septembre 2014.

[24] De mêmes les travaux de recherche, guides méthodologiques et stratégiques, les échanges de bonnes pratiques, etc. se dénombrent par centaines sur le site de l'unité d'évaluation des Nations unies : http://uneval.org/document/library, consulté le 24 septembre 2014.

[25] Les guides méthodologique et papiers de recherche sur l'évaluation se comptent par dizaines sur le site de la Commission européenne : https://ec.europa.eu/europeaid/egeval_en.

[26] Entre autres, ROSSI, Peter H, et al, *Evaluation: A systemic approach*, septième édition, SAGE Publications, 2004, 380p; ARNAUD, Serge, BOUDEVILLE, Nicolas, *Évaluer des politiques et programmes publics*, Paris, Éditions de la Performance, 2004, 136p ; PERRET, Bernard, *L'évaluation des politiques publiques,* Repères, 2008, 120p ; VIVERET, Pierre, *L'évaluation des politiques et des actions publiques*, La Documentation française, 1989.

[27] Entre autres, FURUBO, Jan-Eric, RIST, Ray. C, SANDAHL, Rolf, *International Atlas of Evaluation*, New Brunswick, Transaction Publishers. Cop, 2002, 471p ; *L'évaluation des politiques publiques en Europe: cultures et futurs,* Colloque de Strasbourg, L'Harmattan, 2008, 442p ; JACOB, Steve, *Institutionnaliser l'évaluation des politiques publiques : étude comparée des dispositifs en Belgique, en France, en Suisse et au Pays-Bas*, PIE-Peter Lang, 2006, 271p.

Certains auteurs privilégient davantage une approche de l'évaluation des politiques publiques par ses acteurs, les évaluateurs[28]. Notre recherche se fonde sur leurs travaux dans l'analyse qu'elle fait de l'insertion des évaluateurs dans le système de l'aide au développement. Enfin, nous utilisons les travaux critiques sur l'évaluation qui considèrent cette dernière comme un instrument de « savoir-pouvoir » conduisant à technocratiser la mise en place des programmes de politiques publiques et à renforcer le poids des experts dans la gouvernance des politiques publiques. Selon ces auteurs, l'évaluation renforcerait le fossé entre les représentants de l'État et la population en donnant une image déshumanisée et technique de la prise de décision politique[29]. À partir de la thèse défendue par ces travaux, nous questionnons les relations de pouvoir, ou de domination, qui se créent autour de l'évaluation, ainsi que les effets transformateurs que peuvent avoir les évaluations sur les relations entre les donateurs et les bénéficiaires de l'aide.

Après avoir présenté la littérature existante sur l'évaluation des politiques publiques, et plus spécifiquement sur l'évaluation de l'aide au développement, nous allons désormais introduire le socle théorique et les influences disciplinaires sur lesquels se fonde notre travail de recherche. Nous exposerons également les principaux enjeux qui seront explorés dans les différents chapitres de notre recherche.

- **Enjeux et approches**

Notre recherche réalise une analyse globale de la diffusion d'une culture de l'évaluation dans le secteur de l'aide au développement. Elle raconte l'histoire de la montée de la bonne pratique de l'évaluation dans l'aide au développement, des acteurs qui l'ont portée ou combattue, des enjeux politiques qu'elle soulève et des effets qu'elle a pu avoir sur le régime de l'aide au développement.

La notion de régime, apparue dans un numéro de la revue *International Organisations* de 1982, puis reprise par Stephen Krasner dans son ouvrage *International Regimes*[30], désigne un

[28] Entre autres, PICCIOTTIO, Robert, "The logic of evaluation professionalism", *Evaluation*, vol. 17, n°2, 2011, pp. 165-180 ; MATYJASIK, Nicolas, « Des évaluations de politiques et des consultants: émergence d'un champ professionnel et segmentation des rôles », *Revue française d'administration publique,* n°148, 2010, pp. 907-921 ; KOSHELEVA, Natalia, SEGONE, Marco, « Evalpartners: Facilitating the Development of a New Model of Voluntary Organization for Professional Evaluation to Support the Development of National Evaluation Capacities », *American Journal of Evaluation*, publié en ligne le 30 juillet 2013, pp. 1-6.

[29] Entre autres, ZARKA, Yves-Charles, « L'évaluation, un pouvoir supposé savoir », *Cités*, n°37, janvier 2009, 12p ; del RAY, Angélique, *la Tyrannie de l'évaluation,* La Découverte, Collection Cahiers libres, 2013, 160p ; ELIADIS, Pearl, FURUBO, Jean-Éric, JACOB, Steve, *Evaluation, Seeking Truth or Power ?*, Transaction Publishers, 2011, 358p.

[30] KRASNER, Stephen. D, *International Regimes*, Cornell University Press, 1983, 372p.

« ensemble de principes, de normes, de règles et de procédures de prise de décision implicites ou explicites autour desquels les attentes des acteurs convergent dans un domaine donné »[31]. Un aspect particulièrement intéressant de la théorie de Stephen Krasner est la distinction faite entre, d'une part les principes et les normes qui constituent l'essence du régime, et d'autre part les règles et les procédures qui opérationnalisent ces normes et ces principes. Selon Stephen Krasner, « des changements dans les règles et les procédures de prise de décision sont des changements dans le régime, sous réserve que les normes et principes ne soient pas altérés. Les changements de principes ou de normes sont des changements du régime en lui-même. Quand des normes et des principes sont abandonnés, il y a soit un changement vers un nouveau régime, soit la disparition de régimes dans un domaine donné »[32].

Nous analysons comment la systématisation de l'évaluation a été un changement *dans* le régime de l'aide, en devenant une procédure de prise de décision érigée en bonne pratique. Nous étudions, ensuite, dans quelle mesure l'évaluation a contribué à un changement *du* régime de l'aide, en légitimant de nouveaux principes centrés autour des notions d'efficacité, de redevabilité et de maîtrise des risques des programmes de développement. L'émergence de l'évaluation en tant que bonne pratique routinière a tout d'abord répondu à des contraintes organisationnelles pesant sur les organismes d'aide. En effet, dans le contexte des années 1990 où il fallait relégitimer l'aide face à l'échec des politiques du Consensus de Washington, ces derniers se sont vus obligés par leurs tutelles financières de rendre compte de leurs résultats. Le recours systématique à l'évaluation a renforcé l'importance du résultat, de la transparence, de la preuve et du contrôle, contribuant à la création d'un nouveau régime de l'aide dont le principe central est l'efficacité. Si les bonnes pratiques d'évaluation ont été construites par les donateurs, à mesure que leur respect devient une obligation, ceux-ci se retrouvent pris dans une dynamique de surenchère évaluative qui détermine à son tour leurs politiques et pratiques.

Une fois l'évaluation érigée comme une bonne pratique par les organisations multilatérales, les donateurs bilatéraux se doivent de l'adopter et d'évaluer toujours plus et dans tous les secteurs d'activité. Le succès de l'évaluation provient du fait que son introduction dans le régime de l'aide répond à un besoin fort des donateurs d'avoir des instruments permettant d'améliorer leur redevabilité auprès de leurs tutelles financières et des contribuables, ainsi que de donner une image rationnelle et légitime à une aide au développement perçue comme peu efficace. L'évaluation répond également à un besoin de mémoire organisationnelle sur le

[31] BATISTELLA, Dario, PETITVILLE, Frank, SMOUTS, Marie-Claude, *et al.*, *Dictionnaire des relations internationales : approches, concepts, doctrine*, Dalloz, 2012, 572 p.
[32] Krasner, 1983, *op. cit.* pp. 187-188.

développement et d'accumulation de la connaissance pour ne pas répéter les erreurs du passé et améliorer la gestion des programmes d'aide.

Notre recherche adopte une démarche de sociologie des relations internationales, dont l'objectif consiste à « souligner les continuités et les discontinuités dans les modes d'action, les contraintes et les dynamiques que les acteurs contribuent à créer, mais dans lesquelles ils sont également plus ou moins involontairement pris »[33]. Cette démarche implique « un regard rigoureux sur les réalités internationales comme réalités sociales tant au niveau micro des acteurs qu'au niveau macro des ensembles qu'ils composent »[34]. Nous n'avons pas pour ambition de produire un jugement de valeur sur les différentes méthodes d'évaluation ni sur l'efficacité de l'aide. Notre objectif est d'analyser comment la pratique de l'évaluation a été construite par les acteurs de l'aide pour répondre à des besoins d'apprentissage, de redevabilité, de légitimité, de contrôle, etc.

Nous considérons l'évaluation comme un « instrument de gouvernement » de l'aide au développement. Suivant la distinction proposée par Pierre Lascoumes et Patrick Le Galès, notre analyse de l'évaluation ne porte tant pas sur ses caractéristiques techniques et méthodologiques que sur ses déterminants politiques. À travers le cas d'étude de l'évaluation, nous nous proposons d'analyser « l'instrumentation » de l'aide au développement, c'est-à-dire « l'ensemble des problèmes posés par le choix et l'usage des outils (des techniques, des moyens d'opérer, des dispositifs) qui permettent de matérialiser et d'opérationnaliser l'action gouvernementale »[35]. Il s'agit de révéler les effets politiques que produit l'évaluation sur l'espace socio-politique qu'est l'aide au développement avec ses normes, ses acteurs et ses représentations. C'est donc sous l'angle de l'évaluation comme « instrument de gouvernement » que nous allons analyser le régime actuel de l'aide au développement. Pierre Lascoumes et Patrick Le Galès définissent ce qu'ils appellent « instrument d'action publique » comme « un dispositif à la fois technique et social qui organise les rapports sociaux spécifiques entre la puissance publique et ses destinataires en fonction des représentations et des significations dont il est porteur »[36]. Rejoignant cette approche, nous montrons comment l'évaluation structure les relations et les rapports de force entre les acteurs de l'aide au développement dans leur ensemble, et plus spécifiquement entre les donateurs et les bénéficiaires. Nous rejoignons ainsi les chercheurs qui analysent l'aide au développement d'un point de vue « relationnel » en se

[33] DEVIN, Guillaume, *Sociologie des relations internationales*, La Découverte, 2007, 128p, p. 22.
[34] *Ibid.*
[35] LASCOUMES, Pierre, LE GALES, Patrick, *Gouverner par les instruments*, Presses de Sciences Po, 2005, 369p.
[36] *Ibid.*

concentrant sur les acteurs, leurs logiques d'action, leurs interactions et leurs perceptions[37]. C'est ce lien entre le régime, l'instrument et les acteurs que nous nous proposons d'explorer.

- **Questionnements et hypothèses**

L'idée de cette recherche provient du constat que l'évaluation est omniprésente dans le régime actuel de l'aide au développement. Notre premier questionnement a été de comprendre comment la pratique de l'évaluation est aujourd'hui devenue une bonne pratique incontournable de l'aide au développement. Comment cette bonne pratique de l'évaluation s'est-elle construite ? Quels sont les acteurs qui ont systématisé le recours à l'évaluation ? À quelles logiques organisationnelles et politiques répond cette systématisation dans les organismes de développement ?

Nous nous sommes ensuite interrogée sur les acteurs que sont les évaluateurs et leur insertion dans le système de l'aide au développement. Qui sont les évaluateurs et quel est leur parcours académique et professionnel ? Comment les évaluateurs légitiment-ils leur expertise et se constituent-ils en réseaux professionnels ? Ces réseaux permettent-ils l'émergence d'évaluateurs dans les pays du Sud ayant la capacité de produire un jugement sur l'efficacité, la pertinence et l'utilité des programmes d'aide mis en œuvre par les donateurs ?

Cette dernière interrogation nous a amenée à réfléchir aux effets de l'évaluation sur les relations entre donateurs et bénéficiaires de l'aide. L'évaluation est-elle un instrument permettant aux bénéficiaires de l'aide de mieux s'approprier les programmes d'aide au développement ? L'évaluation permet-elle de rendre plus égalitaires les relations entre donateurs et bénéficiaires en offrant un espace de parole à ces derniers ? L'évaluation n'est-elle pas, au contraire, un instrument renforçant la domination des donateurs sur les bénéficiaires en technicisant l'aide au développement, comme le condamnent les auteurs critiques sur l'évaluation ? Les acteurs de l'aide, donateurs ou bénéficiaires, qui ne sont pas efficaces, sont-ils déclassés, c'est-à-dire se retrouvent-ils en marge du système en raison d'un affaiblissement de leur réputation ?

Nous avons abordé ces trois catégories de questions à partir de plusieurs hypothèses qu'il s'agira ensuite de confirmer ou d'infirmer. Premièrement, nous émettons l'hypothèse que la « bonne pratique » de l'évaluation s'est construite en fonction des évolutions politiques de l'aide

[37] GROVES, Leslie, HINTON, Rachel, *Inclusive aid: changing power and relationships in international development*, Earthscan, 2004, 237p.

au développement au fil des décennies. L'évaluation n'est pas un instrument construit en laboratoire, de manière objective, comme s'élaborerait une théorie mathématique. Elle est une réponse aux contraintes organisationnelles et aux besoins de légitimité et d'apprentissage des acteurs de l'aide en recherche d'une plus grande efficacité de leurs activités. Si l'évaluation peut aujourd'hui déterminer, en partie, l'action des acteurs de l'aide qui ont intégré cette bonne pratique, elle a, en premier lieu, été formatée par les approches politiques dominantes de l'aide au développement. Toutefois, s'il existe aujourd'hui un ensemble de standards internationaux censés harmoniser les pratiques d'évaluation des organismes d'aide, nous pensons que cette uniformisation des pratiques cache des acceptations différentes de l'évaluation par les agences d'aide bilatérales en fonction de leur culture politique nationale.

Notre deuxième hypothèse est que s'il existe aujourd'hui des « bonnes pratiques » de l'évaluation internationalement reconnues, l'expertise des évaluateurs est mal comprise et n'est pas considérée comme légitime par les autres acteurs de l'aide, que ce soit le personnel des agences d'aide ou les gouvernements des pays bénéficiaires. Nous estimons qu'une des raisons de cette situation est que les évaluateurs ne parviennent pas à se constituer en « communauté professionnelle » ou « épistémique » internationale. Le transfert de compétences en évaluation qui a lieu depuis les années 1990 entre les donateurs et les pays bénéficiaires nous apparaît très maigre et les évaluateurs du Nord et du Sud peinent à travailler conjointement sur la base d'un partenariat équilibré. En conséquence, il serait faux de parler d'un réseau transnational d'évaluation coopérant équitablement dans la réalisation des évaluations des programmes d'aide au développement.

Notre troisième hypothèse est que l'évaluation, en tant qu'instrument de jugement de valeur, entraîne des phénomènes de reclassement entre les acteurs de l'aide au développement. Tout d'abord, nous supposons que la production de bonnes pratiques internationales en matière d'évaluation entraîne un clivage entre les donateurs traditionnels et les donateurs émergents, les premiers se conformant à ces bonnes pratiques et étant considérés comme des « bons » donateurs et les seconds les rejetant et étant considérés comme des « mauvais » donateurs. Les donateurs émergents se retrouveraient déclassés sur la scène internationale, mais ce déclassement pourrait jouer en leur faveur au niveau local, dans la mesure où ils apparaissent aux pays bénéficiaires comme des donateurs « alternatifs » et que leur non-respect des « bonnes pratiques » de redevabilité, de suivi-évaluation et de transparence leur permet d'offrir une aide plus souple à ces derniers. Ensuite, nous estimons qu'à mesure que l'évaluation est utilisée pour rationaliser l'allocation de l'aide au développement, les pays les moins « performants » se

trouveraient déclassés dans le système de l'aide. Dans ces pays, l'évaluation serait également utilisée comme un instrument de « gouvernementalité » permettant aux donateurs d'avoir un pouvoir transformateur sur les modes de vie des populations les plus vulnérables et sur la gouvernance des États. Enfin, les donateurs traditionnels étant de plus en plus incités à produire des preuves de résultats quantifiés et soumis à des évaluations par les pairs de leurs performances, nous faisons l'hypothèse qu'ils connaîtraient également un phénomène de reclassement international en faveur de ceux qui développent des activités ayant des résultats massifs, visibles et de court terme.

- **Méthodologie**

Comme expliqué ci-dessus, notre recherche analyse la construction de la bonne pratique de l'évaluation et les effets qu'elle produit sur le régime de l'aide. Nous nous concentrons particulièrement sur les effets normatifs des évaluations sur les relations entre les différentes parties prenantes de l'aide au développement. Les principaux acteurs ayant œuvré à faire de l'évaluation une bonne pratique de l'aide au développement sont les organisations multilatérales, notamment le CAD-OCDE, la Banque mondiale, les Nations unies et la Commission européenne. Notre point de départ a donc été l'analyse de la littérature grise produite par ces organisations au sujet de l'évaluation des années 1960 à nos jours. Nous souhaitions montrer que les standards d'évaluation élaborés au niveau multilatéral ont été appréhendés et mis en application différemment par les agences d'aide bilatérales, selon leur culture nationale de l'évaluation des politiques publiques, leur tradition politique en matière d'aide au développement et leurs contraintes et besoins organisationnels. À cet effet, nous avons choisi deux cas d'études, la France et le Royaume-Uni, car ce sont deux pays qui se caractérisent par des cultures antinomiques de gestion des politiques publiques, notamment en ce qui concerne l'évaluation. Confronter les pratiques de ces deux acteurs nous permet de démontrer que les standards internationaux sont adaptés et traduits de diverses manières au niveau national. Nous avons analysé la littérature grise du *Department for International Development* (DFID) pour le Royaume-Uni et de l'Agence française de développement (AFD), du ministère des affaires étrangères et de la Direction générale Trésor pour la France[38]. Cette analyse documentaire nous a également permis de comparer les évolutions des pratiques d'évaluation

[38] Ces trois structures étant chargées de mettre en œuvre les politiques d'aide au développement de la France.

des organismes d'aide, les avancées de la théorie sur l'évaluation des politiques publiques et les changements des approches dominantes de l'aide au développement.

Nous avons également utilisé les bases de données des organismes d'aide sur les travaux d'évaluation qu'elles mènent pour analyser l'ampleur de la montée en puissance de l'évaluation et de son élargissement à des champs diversifiés de l'aide au développement. L'étude de ces bases de données, nous a également pour mesurer la corrélation entre les résultats des évaluations et des phénomènes tels que la sélectivité des pays bénéficiaires.

Notre recherche intégrant les perceptions des acteurs de l'évaluation, nous avons fondé notre analyse sur une cinquantaine d'entretiens semi-directifs avec des personnes investies dans la promotion d'une culture de l'évaluation dans le secteur de l'aide au développement ou étant confrontées à l'évaluation dans leur travail quotidien. Premièrement, nous avons interrogé des évaluateurs travaillant au sein d'organismes multilatéraux et bilatéraux. Une deuxième catégorie d'acteurs qui nous nous avons interrogés sont les évaluateurs externes travaillant pour le compte de ces organismes. Nous avons également recueilli les témoignages d'évaluateurs des pays du Sud qui travaillent à la construction de systèmes d'évaluation dans leurs pays et/ou qui collaborent avec les organismes d'aide à l'évaluation de programmes de développement. Enfin, nous avons consulté des acteurs de l'aide qui ne sont pas des évaluateurs, afin de comprendre quelle image ils avaient de l'évaluation. Ces entretiens semi-directifs nous ont permis d'avoir des échanges approfondis avec ces acteurs sur leur conception de l'évaluation et sur les enjeux politiques que cette pratique soulève.

Nous avons complété ces entretiens individuels avec l'organisation ou l'observation d'ateliers de travail et de tables rondes en collaboration avec des réseaux d'évaluateurs français et internationaux. Ces groupes de travail nous ont permis de confronter des approches différentes du rôle et de l'utilité de l'évaluation entre les évaluateurs du Nord et du Sud, entre les évaluateurs internes et externes aux agences d'aide et entre des évaluateurs et des « profanes ». Dans le cadre de ces tables rondes, nous avons réuni des personnes d'origines géographiques diverses (Europe, Afrique subsaharienne, Maghreb). Ainsi, notre recherche s'est grandement enrichie de paroles d'acteurs qui incarnent des conceptions contrastées de l'évaluation.

Notre recherche se distingue des travaux académiques habituels sur l'évaluation par son approche principalement empirique, et non théorique, de l'évaluation. Afin de comprendre le travail des évaluateurs, nous avons choisi de nous intégrer à une équipe d'évaluateurs au sein d'une agence d'aide au développement et d'en observer les membres de l'intérieur. Ainsi, nous

avons réalisé un stage de six mois au sein de l'unité d'évaluation de l'AFD, qui s'est poursuivi par plusieurs contrats de consultance. Au total, nous avons réalisé plus d'un an et demi d'observation participante auprès de ces évaluateurs, travaillant avec eux et vivant leur quotidien professionnel. Nous avons pu participer à des travaux d'évaluation et assister aux réunions internes de l'unité d'évaluation et aux réunions avec les autres services de l'agence[39]. Durant cette observation participante, nous avons effectué une mission de terrain de trois mois au Laos pour évaluer les effets environnementaux et sociaux d'un projet de grand barrage financé par l'AFD, la Banque mondiale, le Programme des Nations unies pour le Développement (PNUD) et la Banque européenne d'Investissement. Dans ce contexte, nous avons pu établir un dialogue avec des membres de ces organisations et des fonctionnaires laotiens à propos de l'évaluation des programmes d'aide qu'ils mènent, spécifiquement dans le domaine des grands barrages. Nous avons également pu entrer en contact avec des représentants d'autres organismes d'aide au développement européens et asiatiques présents au Laos. Dans le cadre de l'évaluation du barrage, nous avons participé à de nombreuses réunions entre les évaluateurs de l'AFD, les opérateurs du barrage, les fonctionnaires laotiens et les populations bénéficiaires[40]. Cela nous a permis de comprendre, sur un cas d'étude spécifique, comment l'évaluation était perçue sur le terrain par les bénéficiaires de l'aide et quels jeux de pouvoir se créaient lors des processus d'évaluation parmi les donateurs, ainsi qu'entre les donateurs et l'opérateur de projet. Nous avons également pu observer comment les donateurs bilatéraux, multilatéraux, traditionnels et émergents coopéraient, ou au moins coexistaient, sur le terrain.

Cette observation participante à l'AFD nous a aidée à comprendre comment les évaluateurs œuvrent, quelles sont leurs contraintes et quel rôle ils se donnent. Nous avons pu observer de l'intérieur comment l'évaluation est perçue et utilisée par les acteurs qui ne sont pas évaluateurs.

Nous avons également mené plusieurs autres missions d'évaluation en tant que consultante pour le ministère des affaires étrangères, la Cour des comptes ou des cabinets privés d'évaluation. Ainsi, nous avons pu nous rendre sur d'autres terrains que le Laos, notamment le Maroc et la Côte d'Ivoire, et confronter nos observations faites du Laos à la réalité d'autres pays en développement.

Le danger d'une telle observation participante est de se « confondre » avec son objet d'étude : les évaluateurs et l'expertise qu'ils produisent. Il fut parfois difficile de prendre de la

[39] Voir liste des évènements observés en annexe 9.
[40] Voir liste des évènements observés en annexe 9.

distance par rapport à ces évaluateurs avec lesquels nous avons interagi pendant plus d'un an. Nous avons dépassé cette difficulté en réalisant des entretiens avec d'autres acteurs de l'aide au développement du Nord et du Sud, en côtoyant des chercheurs et en nous retirant du milieu professionnel de l'évaluation à la fin de l'année 2013 pour prendre du recul par rapport à notre expérience.

Parallèlement, nous nous sommes impliquée dans plusieurs réseaux d'évaluateurs : la société française d'évaluation (SFE), le réseau *Evalpartners* du CAD-OCDE et la société européenne d'évaluation. En participant à l'organisation d'évènements sur l'évaluation avec ces structures, nous avons pu rencontrer de manière informelle de nombreux évaluateurs de politiques publiques ou spécialisés dans l'aide au développement[41]. Cela nous a également aidée à comprendre le fonctionnement et les aspirations de ces réseaux d'évaluateurs et le processus de professionnalisation que connaît actuellement la pratique de l'évaluation.

- **Structure de la recherche**

Notre propos sera divisé en trois grandes parties, chacune composée de trois chapitres. La première partie retracera la sociogenèse de la « bonne pratique » de l'évaluation. Notre objectif sera de comprendre les logiques organisationnelles et les dynamiques d'acteurs qui ont permis à l'évaluation de passer du statut de pratique marginale de l'aide au développement à celui de pratique incontournable. Nous découperons cette analyse historique en trois périodes se distinguant de celles choisies par les chercheurs ayant précédemment étudié l'histoire de l'évaluation de l'aide au développement[42].

La première période étudiée s'étendra de 1960 à 1990. Nous montrerons que c'est durant cette période que les « bonnes pratiques internationales » de l'évaluation de l'aide au développement se sont construites. Nous dépeindrons deux grandes phases de l'évaluation, reflétant deux courants politiques de l'aide au développement. Tout d'abord, les années 1960 et 1970, durant lesquelles les organismes d'aide bilatéraux étaient peu soumis à des exigences de transparence et de redevabilité et voyaient peu d'intérêt à évaluer les projets qu'ils mettaient en place. L'évaluation était donc une pratique ponctuelle destinée à mesurer le rapport coût/bénéfices des projets, afin d'en vérifier la rentabilité. Elle était aussi mise en œuvre dans un objectif d'apprentissage, afin de tirer des leçons de projets expérimentaux et de juger de la

[41] Voir liste des évènements observés en annexe 9.
[42] CRACKNELL, 2000, *op. cit.;* REBIEN, 1996, *op. cit.*

pertinence de les généraliser. En effet, l'efficacité de l'aide était peu remise en question et le besoin n'était donc pas ressenti d'évaluer les effets de tous les projets d'aide.

Nous expliquerons que l'année 1979 marqua un changement en raison de la conjonction de plusieurs facteurs. Le choc pétrolier déclencha une crise monétaire internationale entraînant une montée des idéologies néolibérales à travers le monde. Dans ce contexte de crise, les gouvernements, notamment anglo-américains, développèrent un intérêt pour l'évaluation dans l'objectif de rationaliser les dépenses publiques et de mettre fin à des politiques jugées désormais inefficaces. Cet intérêt nouveau pour l'évaluation gagna également l'aide au développement, et sa pratique se généralisa dans les organismes d'aide durant la décennie 1980. Afin de pouvoir comparer les connaissances accumulées dans les agences d'aide grâce à l'évaluation, le CAD-OCDE débuta un travail de normalisation de cette pratique. Cette partie s'arrêtera en 1991, à la veille de l'adoption des Principes du CAD pour l'évaluation.

La deuxième période étudiée débutera en 1991 avec l'adoption par le CAD de ces principes sur l'évaluation, qui vont devenir le référentiel international de « bonnes pratiques » d'évaluation. Nous montrerons que les années 1990 furent la période durant laquelle l'évaluation s'est normée et institutionnalisée dans les organismes de développement selon les principes du CAD-OCDE. C'est également le moment où l'approche néolibérale de l'évaluation servant un objectif de rationalisation de l'aide au développement s'est diffusée à l'ensemble des agences d'aide au développement[43]. Nous montrerons que, durant cette période, l'évaluation a été utilisée pour relégitimer l'aide au développement après l'échec des politiques du Consensus de Washington. Nous observerons alors une indétermination entre une approche qualitative et participative de l'évaluation, servant un objectif que nous qualifions « d'égalitarisation » de l'aide au développement, et une approche quantitative et scientifique, servant un objectif de rationalisation. Ce chapitre s'achèvera en 2000, date de l'adoption des Objectifs du Millénaire pour le Développement (OMD) qui ont fondé un cadre universel pour l'évaluation de l'efficacité de l'aide, centré sur la réduction de la pauvreté et des inégalités.

Dans les années 2000, notre troisième période, les organismes d'aide occidentaux ont tous adopté les « bonnes pratiques » d'évaluation promulguées par le CAD-OCDE et mis en place des réformes pour s'y conformer. Ces « bonnes pratiques » semblent uniformément appliquées dans les organismes de développement et l'aide au développement devient la politique la plus rigoureusement évaluée dans de nombreux États. Nous mettrons en évidence

[43] HIBOU, Béatrice, *La bureaucratisation du monde à l'ère néolibérale,* Éditions la Découverte, 2012, 223p.

que la tendance observée dans les années 1990 d'une évaluation mise en œuvre à des fins de rationalisation s'accentue, à tel point que l'évaluation perd sa fonction d'apprentissage pour devenir un outil gestionnaire de redevabilité. Nous soulignerons également que l'évaluation est mise au service des OMD, qui deviennent le référentiel de l'efficacité de l'aide. Toutes les activités de l'aide et toutes les évaluations doivent contribuer à l'atteinte de ces objectifs. À ce référentiel s'ajoute, en 2005, la Déclaration de Paris sur l'efficacité de l'aide qui établit un ensemble de bonnes pratiques de l'aide au développement, dont le respect sera évalué de manière continue. En plus de cadrer les critères d'évaluation de la « bonne » gestion de l'aide, la Déclaration de Paris édicte des bonnes pratiques supplémentaires en matière d'évaluation. Les pays bénéficiaires, les ONG internationales et les donateurs émergents critiquent la Déclaration de Paris qu'ils estiment avoir été élaborée en cercle fermé, sans que les donateurs occidentaux prennent en compte leurs attentes et approches de la coopération pour le développement. Les donateurs produisent alors un discours mettant en avant une approche plus partenariale de l'aide au développement, qui passe par la mise en œuvre d'évaluations participatives, permettant une meilleure appropriation des programmes d'aide par les pays bénéficiaires. Le devoir d'évaluation devient mutuel : des donateurs vers les bénéficiaires et des bénéficiaires vers les donateurs. Ainsi, chacun serait responsabilisé quant à sa contribution pour le développement.

Nous démontrerons qu'un paradoxe apparaît alors entre, d'une part une pratique néolibérale de l'évaluation par les donateurs qui mettent en œuvre l'évaluation comme un outil de redevabilité envers leurs tutelles financières et les contribuables et, d'autre part la volonté de développer des pratiques plus égalitaires d'évaluation permettant aux pays du Sud de développer des capacités en la matière et de reprendre le contrôle sur leurs politiques de développement. Cette dernière approche de l'évaluation nécessite de créer des compétences en évaluation au Sud. Par ailleurs, les évaluateurs éprouvent des difficultés à concilier la fonction de redevabilité de l'évaluation et celle d'apprentissage. En effet, à mesure que les évaluateurs sont poussés par les dirigeants des agences d'aide à mettre l'accent sur la fonction de redevabilité des évaluations, moins d'efforts sont mis sur l'apprentissage mutuel entre les acteurs du développement et la capitalisation de connaissances, qui sont pourtant des spécificités de l'évaluation par rapport à l'audit et au contrôle interne.

Ces derniers constats nous amènerons à nous questionner, dans une deuxième partie, sur la légitimité actuelle des évaluateurs dans le système de l'aide et sur leurs vecteurs d'influence. Ainsi, après avoir analysé la construction d'une culture de l'évaluation dans le secteur de l'aide

au développement à travers les décennies, nous étudierons comment cette culture se diffuse dans les différents espaces du système de l'aide. À cet effet, nous centrerons notre analyse sur les acteurs que sont les évaluateurs et étudierons comment ils essayent, dans ces espaces, de se légitimer en tant que professionnels du développement et de construire une communauté internationale d'évaluateurs. Notre étude portera sur trois espaces du système de l'aide au développement dans lesquels la culture de l'évaluation se diffuse : les organismes d'aide, les pays bénéficiaires et les réseaux professionnels d'évaluateurs.

Dans un premier chapitre, nous analyserons l'influence qu'a l'évaluateur dans les organismes d'aide au développement. Nous démontrerons que les interactions entre les évaluateurs internes, les évaluateurs externes et les autres personnels des organismes d'aide (dirigeants, managers et opérationnels) ont une influence fondamentale sur la diffusion d'une culture de l'évaluation dans ces enceintes. Nous montrerons que les évaluateurs sont au cœur de jeux d'influence dans les organismes d'aide et que leurs travaux, lorsqu'ils sont porteurs de changement, se confrontent à des résistances fortes. Ainsi, les travaux d'évaluation ne seraient pas que des rapports techniques, mais bien des analyses politiques porteuses de potentiels changements que les autres acteurs tentent de limiter ou d'encourager selon leurs intérêts.

Dans un deuxième chapitre, nous étudierons quels sont les vecteurs mobilisés par les évaluateurs du Nord pour diffuser une culture de l'évaluation dans les pays bénéficiaires de l'aide. Nous démontrerons que les actions de formation et d'assistance technique en matière d'évaluation ont longtemps eu un succès limité, ce qui nous amènera à remettre en cause l'existence d'une expertise d'évaluation solide au Sud qui permettrait la mise en place d'évaluations réellement partenariales entre les donateurs et les bénéficiaires. Nous défendrons la thèse que les pratiques des donateurs lorsqu'ils évaluent leurs programmes d'aide nuisent aux actions de formation qu'ils organisent parallèlement. Nous prouverons également que l'intérêt des experts du Sud pour renforcer leurs compétences en évaluation est limité. Nous montrerons que les systèmes nationaux d'évaluation qui se multiplient dans les pays du Sud sont le plus souvent des coquilles vides en raison de la difficulté de voir émerger une demande réelle d'évaluation dans ces contextes politiques.

Pour dépasser cet état de dissolution de l'expertise de l'évaluation au Sud, nous observerons dans un troisième chapitre que les donateurs ont incité des « champions de l'évaluation » qu'ils ont formés au Sud à mettre en place des associations d'évaluateurs nationales et régionales, afin de créer des communautés de pratique. Nous montrerons que ces associations ont été créées pour permettre une diffusion horizontale, ou « Sud-Sud », de la

culture de l'évaluation rompant avec les transferts de compétences des donateurs vers les bénéficiaires, qui ont été peu efficaces. Nous démontrerons que si ces associations apparaissent être des initiatives d'évaluateurs du Sud, elles sont en fait les fruits des évaluateurs des organisations multilatérales qui ont initié et soutenu leur création. Nous analyserons l'influence qu'ont ces réseaux sur la transformation des politiques publiques des pays du Sud vers une plus grande demande et utilisation de l'évaluation. Nous questionnerons également la capacité de ces réseaux à créer une communauté internationale d'évaluateurs tissant un lien professionnel pérenne entre évaluateurs du Nord et du Sud.

Après avoir montré les opportunités, obstacles et résistances à la diffusion d'une culture de l'évaluation dans le secteur de l'aide au développement, nous explorerons, dans une troisième partie, les effets de reclassement qu'entraîne l'évaluation, aussi bien parmi les donateurs que les bénéficiaires de l'aide. En effet, derrière la professionnalisation des évaluateurs au Nord et la formation d'évaluateurs au Sud, se cache l'enjeu de la légitimation des jugements de valeur produits par l'évaluation sur les politiques et les pratiques des donateurs et des bénéficiaires de l'aide au développement. Se saisir de l'évaluation pour faire valoir sa vision des bonnes pratiques de l'aide au développement et déclasser ceux qui ne s'y conforment pas est, selon nous, un enjeu fort pour les acteurs de l'aide.

Dans un premier chapitre, nous observerons si la dichotomie normative entre les donateurs qui se conforment aux « bonnes pratiques » de l'efficacité de l'aide, et plus spécifiquement de l'évaluation, et les donateurs qui ne s'y conforment pas entraîne un effet de déclassement de ces derniers, qui seraient perçus comme des « mauvais donateurs ». L'évaluation est-elle un enjeu réputationnel pour les donateurs ? À travers les exemples de la Chine et de la Corée du Sud, nous analyserons si des éventuels effets de déclassement s'opèrent effectivement, aussi bien au niveau international du système de l'aide qu'au niveau local, avec les cas d'étude du Laos et du Sénégal.

Le deuxième type de reclassement que nous investiguerons est celui portant sur les bénéficiaires de l'aide. Nous analyserons si l'évaluation entraîne, dans les faits, la dépolitisation et la « rationalisation » des relations entre donateurs et bénéficiaires. Notre objectif sera de démontrer si la mise en œuvre d'évaluations de l'efficacité des programmes d'aide conduit à déclasser les pays bénéficiaires les moins efficaces, qui connaîtraient alors des difficultés à accéder aux fonds d'aide. L'évaluation s'apparenterait alors à une forme de « savoir-pouvoir » des donateurs, légitimant « scientifiquement » leur désignation de qui mérite ou non de recevoir de l'aide au développement. Nous nous demanderons ainsi s'il est vraiment approprié de

considérer l'évaluation comme un outil de domination des donateurs sur les bénéficiaires de l'aide, comme le condamnent certains auteurs critiques, ou s'il est possible d'avoir une lecture plus complexe des rapports de forces qui se nouent entre donateurs et bénéficiaires autour de l'enjeu de l'évaluation de l'efficacité des programmes d'aide.

Enfin, nous analyserons les effets de reclassement que produisent les évaluations des politiques d'aide des donateurs. En effet, les mécanismes « par les pairs » ou indépendants (par des ONG ou des centres de recherche) d'évaluation de l'efficacité organisationnelle et des résultats des donateurs se multiplient depuis plusieurs années. Nous démontrerons que ces évaluations conduisent sont porteuses des approches politiques de l'aide portées par leurs concepteurs. Nous démontrerons qu'elles diffusent une approche de l'aide fondée sur la production de résultats quantitatifs à court termes sur les enjeux considérés comme « prioritaires » dans le cadre des OMD. Nous démontrerons que les classements des organismes de développement produits par ces évaluations entraînent une délégitimation de l'action des organisations onusiennes jugées peu « performantes ». Enfin, nous analyserons si ces évaluations entraînent des réformes dans les organismes de développement qui chercheraient à se conformer aux « bonnes pratiques » qu'elles véhiculent, afin de se reclasser.

PARTIE I : Une histoire politique de l'évaluation de l'aide au développement

Cette première partie a pour objectif d'analyser le processus historique de la diffusion d'une culture de l'évaluation dans le secteur de l'aide au développement[44]. Ce processus est mis en rapport, d'une part avec les évolutions des politiques d'aide au développement, et d'autre part, avec l'avancée des techniques d'évaluation des politiques publiques en général. Les ouvrages retraçant l'histoire de l'évaluation de l'aide au développement sont rares et largement issus des pays anglo-saxons. Nous pouvons citer par exemple, Claus Rebien[45] en 1996, Basil Cracknell[46] en 2000, George Pitman *et al.*[47] en 2005. Ces ouvrages restent cantonnés à une approche très méthodologique et abordent peu la dimension politique de l'évaluation de l'aide[48]. Les liens entre les changements dans les pratiques d'évaluation et les transformations, plus générales, des idéologies de l'aide au développement : « évolutionniste, « tiers-mondiste », « libérale », centrées sur la lutte contre la pauvreté, sont peu analysés. Enfin, ces auteurs offrent une analyse limitée des enjeux politiques qui se créent autour de l'évaluation, en termes de production de normes, de jeu d'acteurs, de stratégies d'influence et de diffusion de nouvelles approches du développement.

Cette première partie s'appuiera sur ces ouvrages et les complétera en développant une vision actualisée de l'histoire politique de l'évaluation de l'aide au développement des années 1960 à aujourd'hui. Elle essaiera de cerner les enjeux politiques qui sous-tendent les évolutions des objets, méthodes et utilisations de l'évaluation. L'évaluation sera présentée comme un outil sociologique et politique plus que comme un outil technique de gestion de l'aide. Nous

[44] Selon le *Dictionnaire de la Science politique* de Guy Hermet *et al* (2001), la culture est : « un système de significations communément partagé par les membres d'une collectivité sociale qui en font l'usage dans leurs interactions (…) La culture est donc instrument de description et d'explication. Elle permet de décrire ce que politique veut dire en un lieu et à un moment donné ». Ainsi, lorsque nous parlons de culture de l'évaluation, et non de pratique d'évaluation ou de technique d'évaluation, nous introduisons l'idée que l'évaluation n'est pas un objet neutre dans la collectivité sociale que serait les acteurs de l'aide au développement, mais un objet porteurs de significations, de valeurs, de normes, qui permet de renforcer ou de contredire des visions politiques de l'aide au développement. Le terme culture souligne également l'importance des interactions qui se nouent autour de l'évaluation qui vont en déterminer la mise en œuvre, mais qui vont également, à leur, tour être influencée par elle.

[45] REBIEN, Claus C., *Evaluating Development Assistance in Theory and in Practice,* Avedbury, 1996, 208p.

[46] CRACKNELL, Basil. E., "Evaluating development assistance: a review of the literature", *Public Administration and Development,* vol. 8, 1988, pp. 75-83.

[47] PITMAN, George. K, FEINSTEIN, Osvaldo. N, INGRAM, Gregory. K, *Evaluating Development Effectiveness,* World Bank Series on Evaluation and Development, n°7, 1997, 373p.

[48] BERLAGE, Lodewijk, STOKKE, Olav, *Evaluating Development Assistance: Approaches and Methods,* EADI Book Series, 1992, 219p; DALE, Reidar, *Evaluating Development Programmes and Projects,* Sage Publication, 1998, 220p.

montrerons que les débats, apparemment techniques, sur les méthodologies de l'évaluation, dissimulent des débats idéologiques sur les fondements de l'aide au développement.

Cette première partie analysera également comment les avancées théoriques et méthodologiques de l'évaluation des politiques publiques ont influencé la diffusion d'une culture de l'évaluation dans les organisations multilatérales et bilatérales de développement. Nous étudierons quels ont été les facteurs internes (évolution des approches de l'aide, des croyances en son efficacité, etc.) et externes (évolution des tendances de gestion des politiques publiques, crises budgétaires nationales, etc.) à l'aide au développement ayant permis une institutionnalisation de l'évaluation dans les organismes d'aide au développement.

Cette première partie montrera pourquoi la culture de l'évaluation s'est diffusée à un rythme très lent jusque dans les années 1980 pour connaître ensuite une croissance exponentielle à partir des années 1990, jusqu'à constituer, dans les années 2000, une bonne pratique internationale à laquelle se conforment tous les donateurs du CAD-OCDE. L'évaluation est devenue un aspect fondamental de la légitimité et de la bonne réputation des agences d'aide, qui développent une pratique intensive de l'évaluation, répondant à des standards internationaux multiples en termes d'indépendance, de méthodologie, de transparence, etc.

Cette recherche tente de décloisonner l'expertise de l'évaluation encore très fragmentée entre les différents champs des politiques publiques. Notre analyse s'appuie sur le croisement de plusieurs disciplines : les relations internationales, les politiques publiques et la sociologie des organisations. Cette pluridisciplinarité permet une compréhension globale du processus de construction d'une norme internationale de l'évaluation. Nous adoptons une démarche de sociologie politique pour aborder l'enjeu de la diffusion de la culture de l'évaluation. Nous étudions ce phénomène à partir des organisations dans lesquelles il se concrétise, des jeux de pouvoir qui l'entourent, et des mécanismes d'apprentissage et de changement qu'elle entraîne auprès des différents niveaux d'acteurs.

Nous considérons l'évaluation comme un outil politique au service des donateurs de l'aide au développement. Les donateurs qui ont utilisé, dans un premier temps, l'évaluation comme un vecteur d'apprentissage et d'amélioration des projets d'aide au développement, l'utilisent de plus en plus comme un moyen de redevabilité et de mesure de la performance. En quelques décennies l'évaluation est passée d'une pratique marginale à une pratique généralisée, dont l'utilité affichée en termes de rationalisation de la prise de décision politique semble

surestimée. Cette partie aura pour objectif de resituer à sa juste place l'évaluation dans le système de l'aide au développement et de comprendre l'ampleur et la nature de son influence.

Cette partie permettra d'établir une typologie des différents modes de diffusion d'une culture de l'évaluation et de mettre en avant la dimension politique de ce processus. Nous observerons que la diffusion d'une pratique et d'une culture de l'évaluation dans les organismes d'aide bilatéraux est le fruit de boucles de rétroaction entre l'influence des politiques publiques nationales et la pression des organisations multilatérales qui ont érigé la pratique de l'évaluation en passage obligé de toute « bonne » action d'aide au développement. Nous étudierons notamment les cas de la France et du Royaume-Uni.

Le premier chapitre couvre la période de 1960 à 1991 et portera sur l'émergence de la pratique de l'évaluation dans le secteur de l'aide au développement. Nous verrons qu'elle est alors influencée par l'approche rationaliste et objectiviste des politiques publiques des pays anglo-américains. De 1960 à 1970, dans un contexte où l'efficacité de l'aide au développement est peu remise en cause, l'évaluation est pratiquée de manière sporadique à des fins d'apprentissage sur les projets d'aide expérimentaux. Avec la crise économique du début des années 1980 et le tournant néolibéral qui l'accompagne, l'évaluation est vue comme un instrument de rationalisation de l'aide et sa pratique prend de l'ampleur, notamment pour contrôler la bonne mise en œuvre des réformes du Consensus de Washington par les pays bénéficiaires de l'aide. Devant la hausse de la pratique de l'évaluation, les bailleurs de fonds réfléchissent à des standards internationaux pour harmoniser les procédures et le contenu des évaluations. Ils sont établis par le CAD-OCDE en 1991 et servent depuis de référentiels pour juger de la qualité des travaux d'évaluation produit par les agences d'aide. C'est là que débute notre deuxième chapitre qui examine l'institutionnalisation de la pratique de l'évaluation dans le secteur de l'aide au développement à partir des nouveaux standards internationaux évoqués ci-dessus. Dans un contexte de « crise » de l'aide au développement, l'évaluation devient un outil de légitimation des politiques des donateurs. La pratique de l'évaluation est alors largement influencée par les principes du New Public Management et se concentre sur l'apport de preuves de résultats quantitatifs permettant aux donateurs de démontrer l'utilité de leurs activités dans les pays en développement. Cette période de crise de l'aide s'achève en 2000 avec l'adoption des Objectifs du Millénaire qui marque un renouveau de l'aide au développement. Les évaluations sont mobilisées pour rendre compte de la contribution des donateurs à l'atteinte de ces objectifs. La préoccupation des bailleurs de fonds pour la mesure de la performance et la redevabilité est toujours grandissante. La Déclaration de Paris, adoptée

en 2005, normalise des bonnes pratiques sur la mise en œuvre de l'aide au développement et donne une place centrale à l'évaluation qui doit répondre à des critères divers (être partenariales avec les pays bénéficiaires, être menées de manière conjointe avec les autres donateurs, apporter des preuves de résultats sur l'action des donateurs, etc.). L'évaluation devient un instrument routinier de la gestion de l'aide au développement. En même temps, nous notons l'émergence du courant de pensée de *l'evidence based policy* qui entraîne une demande d'évaluations expérimentales mesurant les impacts des programmes d'aide afin d'informer la prise de décision par des preuves « scientifiques ». On note ainsi un retour à la vision objectiviste de l'évaluation, qui serait porteuse d'une vérité sur le développement.

Chapitre 1 : 1960-1991, l'évaluation, un outil de rationalisation de l'aide au développement

Dans les années 1960 et 1970, l'évaluation était pratiquée de manière sporadique et peu harmonisée par les organismes d'aide au développement. Cette fonction était perçue comme étant peu utile par les donateurs et peinait à trouver sa place dans les organismes d'aide.

Les années 1979 à 1981 marquèrent un tournant. Le choc pétrolier de 1979 entraîna une crise financière internationale dans un contexte économique déjà difficile. Margaret Thatcher, arrivée à la tête du gouvernement britannique la même année, et Ronald Reagan, accédant à la présidence des États-Unis en 1981, firent du néolibéralisme l'idéologie politico-économique de leur mandat, afin de redresser l'économie de leurs pays. Cette tendance gagna les organismes multilatéraux de développement dominés par l'expertise anglo-américaine. En conséquence, l'aide ne fut plus considérée comme un droit universel, mais comme un droit méritocratique pour les pays qui accepteraient de se réformer à partir des recettes dictées par les donateurs. L'évaluation s'adapta à ce tournant idéologique et servit les besoins de mesure de la performance, du retour sur investissement des projets d'aide et de contrôle des pays bénéficiaires de l'aide.

En 1980, le CAD-OCDE créa un groupe d'experts dédié à la détermination et à la diffusion de règles et de procédures en matière d'évaluation. C'est alors que débuta un travail pour déterminer des standards internationaux d'évaluation qui virent le jour en 1991, marquant le début de l'institutionnalisation de l'évaluation dans les agences d'aide.

A. 1960-1978 : l'évaluation, un outil marginal

De 1960 à 1979, l'évaluation attirait peu l'intérêt des décideurs politiques dans un contexte où l'efficacité de l'aide n'était pas mise en doute et où les budgets d'aide au développement paraissaient illimités. Les organismes d'aide faisaient l'objet de peu de pression en termes de transparence et de redevabilité de la part de la société civile et de leurs tutelles financières. Les quelques évaluations menées étaient empreintes de l'idéologie « évolutionniste » de l'aide au développement populaire à cette époque, fondée sur la croyance que l'importation des modèles économiques du Nord dans les pays du Sud permettrait leur

développement économique. L'évaluation était perçue comme un instrument ayant pour fonction de mesurer le ratio « coût-bénéfice » des projets d'aide ou de permettre un apprentissage sur les projets expérimentaux.

- **L'évaluation, une science « objectiviste » et rationnelle**

C'est aux États-Unis que la discipline de l'évaluation des politiques publiques trouve son origine durant les années 1960 à 1970. L'évaluation émergea et se systématisa comme pratique de gouvernement durant les mandats de Kennedy et de Johnson lors de l'adoption de programmes sociaux de grande ampleur dans le cadre de la politique de « guerre contre la pauvreté » et l'émergence du *Welfare State*[49].

En raison de la forte pression de l'opinion publique et de l'opposition politique demandant des justifications et des informations sur les résultats de ces coûteuses réformes sociales, la pratique de l'évaluation se diffusa comme un outil de légitimation pour le gouvernement. L'évaluation eut donc pour première fonction d'apporter des preuves de réussite, afin de permettre aux politiciens de faire accepter leurs choix politiques et de se montrer transparents.

Dans les années 1960, la diffusion de la pratique de l'évaluation fut tellement importante aux États-Unis qu'elle gagna progressivement le statut de discipline académique en tant que telle. Selon plusieurs chercheurs en sciences sociales[50], l'ouvrage fondateur de la discipline est *Evaluative research : Principles and Practice in Public Service and social action programs,* rédigé en 1967 par Edward Schuman[51]. En 1976, deux sociétés professionnelles d'évaluation furent créées qui fusionnèrent ensuite pour devenir *l'American Evaluation Association*. Cette association est aujourd'hui l'un des plus grands forums d'échange internationaux sur l'évaluation et a largement contribué à la diffusion d'une culture de l'évaluation à travers le monde[52]. Elle fut fondamentale dans la légitimation et le renforcement de la discipline de l'évaluation aux États-Unis, car elle permit le développement d'une réflexion autour de

[49] SPENLEHAUER, Vincent, *L'évaluation des politiques publiques, avatar de la planification,* Thèse pour l'obtention du doctorat de science politique de l'université Pierre Mendes France de Grenoble, 8 décembre 1998, 356p, pp. 88-91.
[50] FURUBO, Jan-Eric, RIST, Ray. C, SANDAHL, Rolf, *International Atlas of Evaluation,* New Brunswick, Transaction Publishers. Cop 2002, 471p, pp. 245-250.
[51] SCHUMAN, Edward. A, *Evaluative Research: Principles and Practice in Public Service and Social Action Programs*, Russel Sage Foundation, 1967, 186p.
[52] Pour une analyse plus détaillée du rôle des « sociétés d'évaluation » dans la diffusion d'une culture de l'évaluation à travers le monde, se référer au chapitre 6.

l'évaluation, la publication d'articles et d'ouvrages scientifiques et une professionnalisation de la discipline.

Toutefois, l'intérêt grandissant pour la pratique de l'évaluation aux États-Unis eut peu d'écho dans les pays européens dans les années 1960 et 1970. Cette pratique resta très limitée, voire inexistante, dans la gestion des politiques publiques des autres États, excepté au Royaume-Uni, en Suède et au Canada. Ces pays avaient en commun de s'être lancés dans des plans nationaux de réformes sociales, d'être gouvernés par des partis réformistes (sociaux-démocrates en Suède, démocrates aux États-Unis et libéraux au Canada) et d'avoir des membres de l'administration souvent issus de formations universitaires en sciences sociales. De plus, ces pays étaient tous dans une situation d'expansion budgétaire de la fonction publique. Ce contexte favorisa une croissance de la pratique de l'évaluation qui apparaissait comme étant l'instrument d'une meilleure planification et comme un élément de réponse au besoin d'information de l'opinion publique et de la classe politique sur les résultats des nouveaux programmes engagés.

Les années 1960 sont souvent présentées comme la « première génération » de l'évaluation[53]. Dans les pays pionniers de l'évaluation, les politiques publiques étaient dominées par une approche rationaliste qui favorisa le succès de cette pratique. En effet, l'approche rationaliste croit en la possibilité d'acquérir un savoir « scientifique » sur les processus sociaux et d'utiliser ce savoir pour les contrôler. Selon ce courant de pensée, l'évaluation pourrait donc être un instrument de rationalisation de la prise de décision, dans la mesure où elle apporte des éléments rigoureux d'expertise permettant d'atteindre l'objectif final que devrait avoir toute politique publique ou organisation sociale : les meilleures efficacité et efficience[54] possibles. Les politiciens et les membres des administrations étant considérés comme des individus rationnels, ils devraient automatiquement se fier aux résultats « scientifiques » des évaluations pour guider leurs décisions.

Nous observons que les premiers théoriciens de l'évaluation, parmi lesquels le sociologue James Coleman[55], le mathématicien et philosophe Michael Scriven[56], ou encore le psychologue Donal Campbell[57], furent largement influencés par une approche positiviste des sciences sociales, qui va de pair avec le rationalisme politique évoqué ci-dessus. Suivant les préceptes

[53] GUBA, Egon. G, LINCOLN, Yvonna S. *Fourth Generation Evaluation*, SAGE Publications, 1989, 294p, pp. 35-45.
[54] Ces deux termes et leur nuance sont expliqués dans le lexique en annexe 7.
[55] COLEMAN, James. S, *Policy Analysis on Major Issues: a compilation of papers prepared for the Commission on the Operation of the Senate*, US Government Printing Office, 1972, 167p.
[56] SCRIVEN, Michael, ROTH, Jane, *Evaluation thesaurus*, Edgepress, 1991, 408p.
[57] CAMPBELL, Donald, *Assessing the impact of planned social change*, paper 8, occasional paper serie, University of Darthmouth, décembre 1976.

d'Auguste Comte ou de John Stuart Mill, ces théoriciens avaient la conviction que les organisations publiques devraient adopter, pour être efficaces, une approche scientifique du gouvernement des sociétés, en remplaçant les « arguments politiques relativistes » par des « preuves scientifiques ». Selon ces sociologues, les sciences sociales ont vocation à établir de manière rigoureuse les lois qui structureraient le fonctionnement des sociétés humaines[58]. Nourris par ce courant, les premiers évaluateurs prônaient l'utilisation de méthodes dites expérimentales provenant du milieu médical pour tester l'efficacité des politiques publiques en comparant des groupes « test » de bénéficiaires de ces politiques à des groupes « témoins » non bénéficiaires. En France et au Royaume-Uni notamment, la sociologie était très marquée par une croyance en les vertus conjuguées de la raison et de la science. La sociologie cherchait à objectiver le social et à sortir du cadre « du concept » pour rentrer dans celui « des faits », se distinguant ainsi de la philosophie. Les faits sociaux transcenderaient les individus et s'imposeraient à eux ; ils deviendraient ainsi, au même titre que les faits naturels, justiciables d'une observation rigoureuse, impersonnelle et quantifiable[59].

Nous observons donc que la discipline de l'évaluation fut forgée dans le cadre d'une approche objectiviste des politiques publiques[60]. Des recherches que nous avons menées, il ressort que les décideurs politiques attendaient des évaluateurs qu'ils soient les « diseurs de vérité » de la société moderne, mesurant objectivement les résultats des politiques publiques pour les guider dans leur prise de décision. Les informations fournies par les évaluations sur les enjeux socio-politiques de l'époque devaient, en conséquence, être mesurables et donc chiffrées. L'art de l'évaluation consistait à transcrire des phénomènes sociaux en des données quantitatives qui permettaient de mesurer leur évolution. Pour se faire entendre des décideurs politiques, les évaluateurs devaient utiliser des méthodes scientifiques et rigoureuses. De manière générale, « la raison statistique » avait un pouvoir légitimateur fort sur les arguments des chercheurs en sciences sociales. L'outil statistique était très apprécié des politiciens, car il permettait de faire ressortir des tendances, des moyennes et des probabilités sur les objets étudiés. Il permettait d'ordonner la société, de la « sécuriser », par rapport au chaos apparent des conduites individuelles, et était considéré comme un moyen de prévoir et de contrôler les dynamiques sociales[61].

[58] COMTE, Auguste, *Discours sur l'ensemble du positivisme,* édition Flammarion, éd. 1999, 1848, 464p.
[59] BOUDOUIN, Jean, *Introduction à la sociologie politique,* Éditions du Seuil, 1998, 324p.
[60] REBIEN, 1996, *op. cit.* pp. 50-55.
[61] DEROSIERES, Alain, *La politique des grands nombres, essai sur la raison statistique*, Paris, La Découverte, 2000, 456 p.

Dans les années 1960, l'évaluation était donc considérée comme un outil de « savoir-pouvoir » à destination des décideurs politiques. Elle servait à diminuer les risques déstabilisateurs des réformes expérimentées. À cette époque, les évaluateurs étaient principalement des experts en sciences sociales ayant une très bonne connaissance du contenu des politiques évaluées, mais une faible expertise en matière de sociologie des organisations et politique. En conséquence, les évaluations se concentraient sur l'étude des réalisations des politiques publiques sans aborder leurs processus de mise en œuvre et les enjeux de gouvernance. De plus, les évaluations se focalisaient sur des groupes très ciblés de bénéficiaires sans regarder les impacts[62] plus globaux des politiques publiques. Elles abordaient les politiques publiques comme des « boîtes noires »[63], dont l'ouverture serait inutile puisque les objectifs et les résultats seraient liés par des chaînes de causalité linéaires pouvant être déterminées *ex ante*.

L'approche rationaliste impliquait un usage direct des évaluations dans la prise de décision. En théorie, le décideur politique rationnel devrait prendre directement en compte les preuves scientifiques rigoureuses des évaluateurs, puisque son intérêt est l'amélioration de l'efficacité et de l'efficience de sa politique. Les évaluations étaient censées faire des recommandations directement applicables par le décideur politique pour répondre aux problématiques qui se posaient à lui. Les évaluateurs ne se sentaient donc pas concernés par la manière dont les évaluations seraient utilisées, d'une part car leur utilisation était théoriquement automatique, d'autre part car cela nuirait à leur objectivité.

Les évaluations étaient donc principalement réalisées dans un objectif « sommatif »[64] et selon un format expérimental. Donald Campbell estime que « la société et la nation devraient être vues comme des laboratoires d'expérimentation sociale. Les évaluations devraient appliquer des méthodologies expérimentales pour répondre à des questions sommatives et fournir des conclusions claires facilement utilisables par les politiciens. Les expérimentations aléatoires pourraient ainsi aider à identifier les meilleures interventions sociales, afin qu'elles soient ensuite généralisées »[65]. Ainsi, les évaluations permettraient de décrypter les chaînes de causalité des politiques publiques pour en comprendre les résultats et permettre de les améliorer.

[62] La définition du terme « impact » en matière d'évaluation est présentée dans le lexique en annexe 7.
[63] Terme utilisé en science politique pour illustrer l'impression d'opacité que donnent les politiques publiques lorsqu'on les analyse de l'extérieur en ne s'intéressant qu'à leurs effets et non à leur processus de fabrication et de mise en œuvre.
[64] Les évaluations sommatives sont des évaluations mises en place dans le but de faire le bilan des résultats apportés par une action publique et de décider en conséquence de son maintien ou de son retrait. Elles sont mises en place dans un objectif de redevabilité et utilisées par le décideur pour démontrer que les politiques qu'il met en œuvre sont efficaces.
[65] CAMPBELL, Donald. T, "Reforms as experiments", *American Psychologist*, n° 24, 1969, pp. 409-419.

L'évaluation était également vue comme un outil d'apprentissage lorsque des réformes innovantes étaient mises en place.

Le débat le plus important parmi les évaluateurs portait sur le fait de savoir si les évaluations devaient émettre ou non des jugements de valeur sur les politiques qu'elles traitaient. Certains évaluateurs, comme Michael Scriven[66], considéraient le jugement de l'évaluateur comme une condition primordiale de la prise en compte des résultats des évaluations par les décideurs politiques. L'évaluateur devait servir de guide à l'homme politique. Pour d'autres[67], au contraire, les évaluations devaient rester en dehors du politique pour préserver leur objectivité. Il revenait aux décideurs politiques d'interpréter et de s'approprier les recommandations formulées. Dans un contexte positiviste des sciences sociales, c'est la deuxième approche qui l'emporta. En effet, selon ce courant, le chercheur ou l'expert en sociologie politique doit s'auto-distancier par rapport à la société qu'il observe et éliminer tout jugement de valeur. C'est cette capacité à prendre du recul sur la société qui fait sa supériorité et la valeur ajoutée de son expertise par rapport aux autres citoyens. Il doit poser un regard scientifique, froid et clinique sur la société[68]. L'évaluateur, pour être crédible et accepté, devait donc proscrire tout jugement de valeur qui serait immédiatement jugé impressionniste et subjectif.

Aux États-Unis, l'approche rationaliste des sciences sociales et de l'évaluation fut incarnée par la mise en place, en 1965, par le président Johnson, du *Program, Planning and Budgeting System* (PPBS) d'abord appliqué au secteur militaire, puis étendu à l'ensemble des politiques publiques. Cette méthode de gestion administrative imbriquait les phases de programmation, de planification, de budgétisation et d'évaluation des politiques publiques dans un tout cohérent. Ces phases devaient être rationnellement liées par des modélisations économétriques. Aux objectifs programmés correspondaient une planification et un budget qui devaient aboutir à des résultats qui seraient ensuite évalués. L'évaluation servait alors de base de départ à un nouveau cycle de programmation. Vincent Spenlehauer explique que le PPBS a « permis d'ancrer pan-institutionnellement la notion et la pratique systématique de l'évaluation de programmes dans l'administration fédérale américaine »[69]. Les décideurs politiques l'utilisaient pour donner une justification scientifique et rationnelle aux dépenses publiques

[66] SCRIVEN, Michael, "The Methodology of evaluation, in TYLER, Ralph, GAGNE, Robert, and SCRIVEN, Michael, *Perspective of curriculum evaluation,* Rand McNally, 1967, pp. 39-83.
[67] CAMPBELL, Donald. T, 1969, *op. cit.*
[68] DURKHEIM, Émile, *Les règles de la méthode sociologique*, Paris: Les Presses universitaires de France, 16e édition, 1967, première publication en 1894, 149 p.
[69] SPENLEHAUER, Vincent, *L'évaluation des politiques publiques, avatar de la planification,* Thèse pour l'obtention du doctorat de science politique de l'université Pierre Mendes France de Grenoble, 8 décembre 1998, 356p, p. 91.

qu'ils engageaient. Le PPBS constituait « une forme de macro-technologie gouvernementale conçue pour comparer et hiérarchiser tous les programmes publics quels qu'ils soient »[70]. Les évaluations des programmes sociaux référaient donc à une technicité économiste de type calcul « coût/bénéfice ». Nous devinons ici que les politiciens cherchaient à se protéger derrière une pratique très technocratique de l'évaluation, dont le vocabulaire était austère et seulement compréhensible par un nombre limité d'initiés. Techniciser la mise en œuvre des politiques publiques permettait de dépolitiser les enjeux sociétaux traités et de contrer les attaques « molles » de la société civile par les preuves « dures » des évaluations scientifiques.

Toutefois, dès la fin des années 1960, les experts en sciences sociales américains commencèrent à s'éloigner de l'analyse causale et rationnelle des phénomènes politiques et sociaux et à s'intéresser davantage à la complexité des processus de changement dans la société. Nous retrouvons cette tendance dans le milieu de l'évaluation, où les théoriciens contestèrent de plus en plus la vision de la réalité décrite jusqu'alors comme une chaîne linéaire de causes à effets[71]. Le sociologue Mancur Olson, ou encore le politologue Walter Mondale, s'élevèrent contre les méthodes microéconomiques des évaluations du PPBS qu'ils considéraient peu adaptées aux programmes sociaux. Toutefois, ils conservèrent dans leur approche les principes d'indépendance et d'objectivité des évaluations issus de la doctrine du PPBS. Le gouvernement des États-Unis commandita un rapport à Mancur Olson, afin qu'il propose des méthodes alternatives de gestion et d'évaluation des politiques sociales. Dans ce rapport, il déclara « vouloir développer des types différents d'évaluation des politiques publiques : quasi-expérimentation sociale, évaluation participative, analyse qualitative des politiques publiques (reconstitution des théories d'action, analyse de mise en œuvre, etc.) »[72]. Il estimait, en effet, que « le PPBS et les autres méthodes modernes, utilisés pour résoudre les problèmes politiques, qui mettent l'accent sur l'intérêt d'une quantification rigoureuse, donnent trop d'importance aux grandeurs mesurables au détriment de celles qui ne le sont pas (…) Une meilleure évaluation des conditions sociales de la population et une détermination plus équilibrée des priorités sociales, exigent de disposer d'instruments de mesure plus complets »[73]. Ainsi, les sociologues américains commencèrent déjà, dans les années 1970, à développer des approches

[70] *Ibid*.
[71] WEISS, Carol H., "The stakeholder approach to evaluation: Origins and promise", *New Directions in Program Evaluation*, n°17, 1983, pp. 3-14.
[72] OLSON, Mancur, *Toward a social report*, US Department of Health, Education, and Welfare. Washington: U.S Government Printing Office, 1969.
[73] OLSON, Mancur, 1969, *op. cit.*

« sociologiques » de l'évaluation, utilisant des instruments d'analyse qualitative, comme les enquêtes, les entretiens ou les *focus groups*.

Nous voyons ici que les débats de méthode sur la gestion et l'évaluation des programmes sociaux reflétaient des débats politiques plus profonds sur la manière de « bien gouverner ». Nous avons, d'une part une approche qui privilégiait l'efficience économique et financière des politiques publiques et l'évaluation expérimentale et, d'autre part une tendance qui privilégiait la recherche de l'équité et du bien-être social et favorisait des méthodes d'évaluation qualitatives et participatives. Toutefois, cette première percée de l'approche sociologique qualitative n'a pas survécu au changement de gouvernement, le président Nixon ayant laissé l'élaboration d'une comptabilité sociale fédérale à la charge des instituts statistiques, largement issus d'une culture économiste. Néanmoins, selon Vincent Spenlehauer, « grâce au tapage de la controverse scientifico-gouvernementale à propos des indicateurs sociaux, l'évaluation des programmes de politiques publiques, dont le principe a été institutionnalisé dans la foulée du lancement du PPBS, a évolué très sensiblement. L'évaluation s'est enrichie de considérations sociales. La notion d'efficacité sociale *ex post* des programmes publics s'est imposée à égalité aux côtés de la notion d'efficience économico-budgétaire pratiquée par les chargés d'études du PPBS »[74].

En plus de ces débats, après une décennie de pratique de l'évaluation, il apparut que les recommandations des évaluations étaient loin d'avoir été directement suivies de prises de décision[75]. Les évaluateurs conceptualisèrent alors une approche alternative de l'utilisation des évaluations, qualifiée d'« incrémentale », ou de « conceptuelle », selon laquelle elles n'auraient pas une utilité directe, mais diffuse et sur le long terme[76]. L'évaluation ferait partie d'un faisceau d'informations dont dispose le décideur pour gouverner. C'est l'arbitrage entre ces informations, dont l'évaluation fait partie, qui entraînerait un changement de politiques publiques, et non une évaluation isolée. Selon cette approche, les décideurs politiques n'étaient plus perçus comme des idéaux-types rationnels. De plus, les transformations radicales de politiques étaient considérées comme des exceptions, même lorsque leur nécessité était prouvée de manière « scientifique ». Il serait alors normal que l'évaluation n'ait pas un usage direct qui se manifeste par des changements très visibles. Les évaluations seraient utilisées sur le long terme, indirectement, par l'accumulation des connaissances qu'elles produisent. Leur usage premier serait de transformer les perceptions des acteurs des politiques publiques, ainsi que leur

[74] SPENLEHAUER, Vincent, *op. cit.* p. 110.
[75] WEISS, Carol, 1883, *art. cit.*
[76] PATTON, Michael, *Utilization focused evaluations,* Sage Publication, 1978, 688p, pp. 100-101.

manière de penser et de conduire leurs activités, et non d'entraîner des réorientations majeures des politiques publiques.

Ce nouveau courant de l'évaluation donna également plus de place aux acteurs dans les processus d'évaluation. Le terme de *stakeholder* (partie prenante) fut pour la première fois appliqué à la discipline de l'évaluation dans le milieu des années 1970 pour introduire l'idée que tous les groupes d'acteurs participant à une politique publique devraient être impliqués dans son évaluation[77]. Certains évaluateurs prirent conscience que l'application des recommandations ne se faisait pas de manière automatique, mais demandait une appropriation des recommandations par les acteurs concernés, quelle que soit la scientificité et la justesse des preuves apportées par l'évaluation[78]. Ils eurent, dès lors, pour objectif d'apporter des réponses aux besoins d'information des différents groupes d'acteurs. Ils ouvrirent la « boîte noire » des politiques publiques et commencèrent à s'intéresser à l'analyse de leurs processus de fabrication et de mise en œuvre, en considérant ces processus comme non déterminés et influençant la réussite des programmes évalués. À cette époque, les évaluateurs statuèrent clairement que leur rôle n'était pas de porter un jugement sur les politiques publiques. Les parties prenantes devaient construire leur propre jugement à partir des observations faites par les évaluateurs qui n'avaient pas de pouvoir d'autorité.

Nous retenons de cette analyse que pour que l'évaluation devienne une bonne pratique dans les administrations publiques, il faut un consensus tripartite sur son importance entre la communauté scientifique (les chercheurs), les opérateurs des politiques publiques (l'administration) et le pouvoir politique (les gouvernants). Le maintien de la légitimité des méthodes prédominantes d'évaluation dans ces trois cercles est une condition de leur pérennité. Un cercle vertueux s'installe alors, au sein duquel ces différents acteurs renforcent et légitiment mutuellement leur argumentaire pour faire d'une approche spécifique de l'évaluation et des politiques publiques une norme.

Nous postulions, en début de chapitre, que la pratique de l'évaluation ne s'était pas diffusée en dehors des pays anglo-saxons dans les années 1960 et 1970. Cette affirmation peut paraître surprenante, dans la mesure où il est commun de voir en la politique de « rationalisation des choix budgétaires » (RCB), développée par l'administration française durant cette période, un avatar du PPBS des États-Unis. Nous estimons que la RCB trouve davantage ses origines dans une culture colbertiste de l'administration et dans un un courant keynésien de l'analyse

[77] REBIEN, 1996, *op. cit.* pp. 47-49.
[78] PATTON, Michael, 1978, *op. cit.*

économique, qui conduisaient à considérer toute économie comme un système matriciel dans lequel il fallait trouver des multiplicateurs. Si la RCB provenait d'une même volonté de rationalisation des dépenses publiques, elle n'avait pas les mêmes caractéristiques de fonctionnement que le PPBS, notamment en matière d'évaluation[79]. Si les fondateurs de la RCB s'inspirèrent du PPBS, ils avaient conscience que le contexte français était très différent du contexte des États-Unis. De plus, le PPBS était déjà en perte de vitesse lors de la conception de la RCB, ce qui conduisit l'administration française à prendre du recul face à ce modèle.

La RCB reposait bien sur une approche rationnelle, scientifique et centralisée de la gestion publique, fondée sur le calcul de la rentabilité économique de l'action publique. Toutefois, la RCB n'a pas permis d'ancrer en France une culture de l'évaluation *ex post*. En effet, la rationalisation des politiques publiques portait sur la phase précédant la prise de décision politique, ce qui inscrivit l'évaluation dans une logique purement *ex ante*. L'évaluation était perçue comme un outil d'analyse de la faisabilité des politiques publiques, mais pas comme un instrument de gestion ou de mesure des résultats des politiques publiques. Ce ne fut qu'à partir des années 1970 que le milieu de la recherche économique et statistique français s'intéressa aux évaluations *ex post,* lorsque l'approche de la RCB fut largement remise en cause. Ce ne fut qu'après1980 que l'évaluation *ex post* se diffusa dans la gestion des politiques publiques en France. Jusque-là, la thématique de l'évaluation n'avait pas suscité de débats tripartites entre le milieu de la recherche, les opérateurs des politiques publiques et les gouvernants. Les tenants d'une institutionnalisation de l'évaluation dans ces trois cercles étaient peu nombreux, ce qui avait empêché une mise sur l'agenda politique de l'évaluation, comme cela avait été le cas aux États-Unis[80].

Nous voyons donc que la fonction de l'évaluation et son utilisation sont déterminées par le contexte scientifique et sociologique dans lequel les politiques publiques évaluées s'inscrivent. L'évaluation n'est pas un simple instrument de gestion qu'il serait possible d'importer dans un système de politiques publiques sans en bouleverser la culture et la structuration. C'est en ce sens que nous parlons de diffusion d'une « culture » de l'évaluation et pas seulement d'une « pratique » de l'évaluation. C'est également la raison pour laquelle nous faisons l'hypothèse que la culture de l'évaluation a tendance à se diffuser de manière

[79] SPENLEHAUER, Vincent, 1998, *op. cit.*

[80] La mise sur agenda est un concept de sociologie politique théorisé par John Kingdom dans son ouvrage *Agendas, alternatives, and public policies.* Elle désigne « l'ensemble des problèmes faisant l'objet d'un traitement, sous quelque forme que ce soit, de la part des autorités publiques et donc susceptibles de faire l'objet d'une ou plusieurs décisions ». Il faut également souligner que « la mise à l'agenda appelle et justifie une intervention publique légitime sous la forme d'une décision des autorités publiques, quelles qu'en soient la forme (gouvernementale, législative, administrative) et la modalité (mesures d'urgence, plan d'action à plus long terme, loi, décision administrative ou financière, désignation d'une commission, etc.) ».

sectorielle *via* des réseaux académiques, politiques ou professionnels partageant les mêmes valeurs, plutôt que *via* des transferts de connaissances et de pratiques interministériels.

- **Une culture de l'évaluation cantonnée à quelques agences d'aide**

Nous trouvons les premières tentatives de théorisation de l'évaluation de l'aide au développement au début des années 1950. Tout, d'abord dans un article de *l'International Social Science Bulletin* de 1955 : « l'évaluation est un effort d'apprendre ce qui entraîne le changement pendant et après un programme d'action. Ce processus d'apprentissage aura davantage de chances de réussir si l'évaluation est planifiée depuis le début, au même moment que le programme lui-même. La question pratique majeure que l'évaluation doit traiter est dans quelle mesure le programme a atteint ses objectifs ? Il est impossible de répondre à cette question de manière satisfaisante si les objectifs ne sont pas clairement définis »[81]. Ensuite, dans un rapport de Samuel Hayes pour l'UNESCO, dans lequel il définit l'évaluation comme « un ensemble de concepts et de pratiques appliqué de manière à contribuer à l'amélioration des activités pratiques d'aide au développement »[82].

Les premières expériences d'évaluation de l'aide au développement eurent lieu dans les organisations d'aide dominées par l'expertise des États-Unis : la Banque mondiale, les Nations unies, l'OCDE et USAID. La lecture de documents internes à ces organisations sur l'histoire de l'évaluation et les entretiens que nous avons menés auprès d'anciens évaluateurs de la Banque mondiale et des Nations unies[83] font, en effet, ressortir une forte pression américaine pour que ces organisations développent des capacités en évaluation[84].

L'évaluation émergea donc dans le secteur de l'aide en s'inspirant de l'approche de l'évaluation des États-Unis, fortement marquée par un modèle rationaliste et objectiviste des politiques publiques. Nous observons ici un transfert de la bonne pratique de l'évaluation du niveau national (administrations états-uniennes, dont l'agence d'aide USAID) vers le niveau multilatéral (Banque mondiale, OCDE, Nations unies, etc.). Pour comprendre cette émergence

[81] JAHODA, Marie, BARITZ, Edward, « The nature of evaluation », *International Social Science Bulletin*, vol. 7, n°3, 1955, pp. 1-13, p. 1-2.
[82] HAYES, Samuel, *Evaluating Development Projects*, UNESCO, 1959, p.16.
[83] Entretien avec une ancienne évaluatrice de l'OED, Banque mondiale, 20 novembre 2013 ; entretien avec le responsable du réseau d'évaluateurs du CAD-OCDE, 16 novembre 2012 ; archives « Oral History » de la Banque mondiale.
[84] Mc Namara, nommé directeur général de la Banque mondiale en 1971 a été le concepteur de la méthode d'analyse des politiques publiques Planning, Programming and Budgeting System (PPBS) au sein du ministère de la Défense américain dans les années 1960. A la tête de la Banque mondiale, il a donc encouragé le développement de la pratique de l'évaluation selon ce modèle au sein de l'organisation sous la recommandation du gouvernement américain.

de l'évaluation dans les organismes d'aide multilatéraux, il faut la replacer dans le contexte général de la politique de coopération internationale pour le développement dans les années 1960.

En 1962, le CAD-OCDE entreprit de mener des examens réguliers des politiques de développement de ses États membres[85]. Ces examens se restreignaient largement à la mesure des montants d'aide alloués aux pays en développement. Ces montants étaient évalués à partir du standard établi par les Nations unies (au sein de la CNUCED), selon lequel les donateurs s'engageaient à consacrer 1 % de leur PNB à l'assistance aux pays en développement. L'année suivante, l'OCDE, conjointement avec la BIRD de la Banque mondiale, élaborèrent des bases de données sur les montants d'aide fournis par les donateurs bilatéraux et multilatéraux[86]. Ces bases de données permettaient de comparer les allocations des donateurs et devaient inciter ces derniers à être transparents et à améliorer leurs pratiques pour apparaître comme de « bons élèves ». Nous observons ici le début de la culture de la transparence et de la redevabilité au sein de la « communauté des donateurs »[87], qui préexista à et conditionna l'émergence d'une culture de l'évaluation.

Ces exercices de mesure se réduisaient aux volumes d'aide. Les donateurs n'avaient pas d'obligation de redevabilité sur la manière dont ils mettaient en œuvre leurs projets d'aide. L'essor d'une évaluation qualitative du « contenu » des politiques d'aide au développement fut notamment empêché par le flou qui entourait la définition de l'aide au développement jusqu'à la fin des années 1960. Ce n'est qu'en 1969 que l'APD fut officiellement reconnue comme un élément distinct des autres transferts officiels internationaux, comme par exemple les investissements directs à l'étranger[88]. Une définition formelle de l'APD fut adoptée par le CAD en 1972[89]. Elle établissait que les programmes d'aide devaient être composés à 84 % de subventions (dons). En 1986, cette cible fut élevée à 86 %.

La fin des années 1960 et le début des années 1970 furent donc marqués par le développement d'indicateurs et d'instruments pour « évaluer » quantitativement l'allocation

[85] Le CAD fut créé le 5 octobre 1961 en remplacement du Groupe sur l'aide au développement (GAD) établi l'année précédente.
[86] *L'Expanded Reporting System on External Lending* pour la Banque mondiale et le *Creditor Reporting System* pour l'OCDE.
[87] Par « communauté internationale du développement », nous entendons l'ensemble des acteurs qui produisent des politiques et des actions de développement. Elle regroupe donc : les organismes d'aide bilatéraux et multilatéraux, les ONG internationales et nationales, les chercheurs, les États bénéficiaires de l'aide, les populations bénéficiaires de l'aide et l'opinion publique internationale mobilisée sur les enjeux de développement. L'expression est employée avec précaution car cette « communauté » n'est pas formalisée. Elle est utilisée pour faire transparaitre l'émergence d'un consensus « global » sur certains enjeux du développement, notamment ce qui constitue les bonnes pratiques de la gestion de l'aide et l'efficacité de l'aide. Elle n'exclut pas l'existence de contestations de cette opinion commune par des acteurs qui défendent une pensée du développement alternative.
[88] STOKKE, Olav, *The UN and Development, from aid to cooperation*, Indiana University Press, 2009, 721p, p. 170.
[89] DAC-OECD, *Investir dans le Tiers monde*, 1972.

d'aide des pays donateurs. Les organismes multilatéraux d'aide au développement et leurs membres anglo-américains poussèrent les donateurs bilatéraux à être davantage transparents, lesquels répondirent plus (donateurs de l'Europe du Nord) ou moins (France, Allemagne) favorablement, selon l'importance donnée à la transparence dans leurs politiques publiques nationales. Si le terme d'évaluation est utilisé, nous estimons, toutefois, que la pratique des organismes multilatéraux se situait alors davantage au niveau de la mesure que de l'évaluation[90].

Dans les années 1960, la pratique de l'évaluation s'était donc peu diffusée des organisations multilatérales de développement vers les agences bilatérales. La plupart commençaient à pratiquer l'évaluation, mais de manière informelle et ponctuelle. Les évaluations étaient surtout mises en œuvre sur des projets expérimentaux, desquels les donateurs estimaient avoir besoin de tirer des leçons pour décider de leur généralisation. En conséquence, il y avait très peu d'expertise disponible en évaluation et pratiquement aucune ressource n'était allouée à cette activité. La plupart des agences d'aide n'avaient pas de service dédié à cette fonction. Un agent de l'*Oversea Development Administration* (ODA)[91] au Royaume-Uni raconte que « *dans les années 1970, lorsque je travaillais à l'ODA, j'étais forcé d'utiliser des étudiants pour mener des évaluations car il n'y avait ni ressources financières pour engager des consultants, ni humaines pour mobiliser à cette tâche des experts confirmés en interne* »[92]. Ce témoignage illustre le faible intérêt porté à l'évaluation par les dirigeants des agences d'aide. Nous expliquons cette situation par le fait que la culture de l'évaluation des politiques publique n'était pas encore présente dans la grande majorité des pays donateurs. Toutefois, même dans les pays où cette culture était la plus développée (États-Unis, Royaume-Uni, Danemark, etc.), l'aide au développement était peu évaluée par rapport aux autres politiques publiques. Tendance qui s'inversera au fil du temps.

De manière générale, la pensée politique de l'aide au développement des années 1960 et 1970 eut un effet de blocage sur la diffusion d'une culture de l'évaluation dans l'aide au développement.

Tout d'abord, les années 1960 furent marquées par un contexte d'optimisme des acteurs du développement. En juillet 1960, l'Assemblée générale des Nations unies (AGNU) proclama la Décennie pour le développement, durant laquelle les États du Nord et du Sud devaient

[90] Pour une explication de la distinction entre « mesure » et « évaluation » voir introduction.
[91] Ancien nom du ministère de la coopération britannique.
[92] Entretien avec un ancien évaluateur interne de l'ODA, 14 janvier 2013.

s'engager à « intensifier leurs efforts pour mobiliser et maintenir les moyens nécessaires à l'accélération des progrès vers une croissance durable des nations et leur avancement social »[93]. Les objectifs que se fixèrent les États étaient très ambitieux : un taux de croissance annuel de 5 % du PIB des pays en développement et, comme nous l'avons évoqué, 1 % du PNB des pays du Nord consacré à l'aide au développement d'ici à la fin de la décennie. Ces engagements témoignaient d'une forte croyance en l'efficacité de l'aide pour diminuer les écarts de développement entre le Nord et le Sud. Les donateurs étaient dans le temps des engagements, pas de l'évaluation.

Néanmoins, au sein des Nations unies, les représentants des États-Unis poussèrent pour que cette augmentation des fonds d'APD s'accompagne de dispositifs de suivi et d'évaluation des résultats. Leur plaidoyer conduisit la Déclaration de la Décennie pour le Développement à souligner que « le besoin de nouveaux moyens de collecte, d'agrégation, d'analyse et de diffusion de statistiques et autres informations, requis pour suivre le développement économique et social et pour fournir une mesure constante des progrès vers les objectifs de la décennie »[94]. Cependant, ces outils de mesure étaient conçus comme des instruments « d'incitation » à destination des gouvernements des pays du Nord pour qu'ils s'engagent plus généreusement dans l'aide au développement, et pas comme des instruments « de conviction » nécessitant la construction de preuves rigoureuses de résultats pour répondre à une demande de redevabilité. Les donateurs n'avaient pas besoin de ces preuves, puisque l'éthique et l'efficacité de l'aide n'étaient pas remises en question.

Les évaluations n'étaient donc pas considérées comme un exercice nécessaire, ni en termes d'apprentissage (hormis les cas de projets expérimentaux), ni en termes de redevabilité. Cette forte confiance en l'efficacité de l'aide au développement provenait d'une approche mécaniste et universaliste des processus de développement, qui fut notamment théorisée et diffusée par Rostow[95]. Selon lui, le développement des États est déterminé par leur croissance économique et suit cinq étapes : la société traditionnelle, les conditions préalables du décollage économique, le démarrage, le progrès vers la maturité et la consommation de masse. Il modélise les économies des pays non industrialisés à partir de celles des États occidentaux, estimant qu'elles devraient suivre le même processus de développement. Il faut rappeler que cette théorie fut développée dans un contexte de guerre froide et servait à prouver la supériorité du modèle

[93] Dans STOKKE, Olav, *The UN and Development, from aid to cooperation*, Indiana University Press, 2009, 721 p, pp. 139-141.
[94] Assemblée générale des Nations Unies, *Déclaration de la Décennie pour le Développement*, 1960.
[95] ROSTOW, William, *The Stages of Economic Growth: A Non-Communist Manifesto*, Cambridge University Press, 1960, third ed. 1991, 324p.

économique capitaliste sur le modèle communiste pour atteindre le développement. Cette dimension stratégique contribua sûrement à la préférence portée par les politiciens à cette approche pendant plusieurs décennies. Elle fut reprise par de nombreux chercheurs en science politique, dont Gabriel Almond et James Coleman[96], selon qui tous les pays en développement seraient automatiquement gagnés par la modernisation politique que connaissent les États occidentaux s'ils en suivaient le modèle.

En reprenant le terme de Bertrand Badie, nous identifions cette pratique des bailleurs de fonds à une nouvelle tentative « d'importation de l'État occidental »[97]. Les pays donateurs essayaient de modeler les États du Sud à l'image des États du Nord, dans une logique selon laquelle la clé du développement des pays sous-développés était la reproduction du modèle de développement des États occidentaux. Selon ces approches « évolutionnistes » des processus de développement, l'évaluation de l'efficacité de l'aide au développement est un exercice inutile puisque les États du Sud parviendront de manière linéaire au développement sous la tutelle des donateurs. Cette vision du développement permettait également de justifier les politiques coloniales, à une période où elles étaient remises en cause et où les mouvements indépendantistes prenaient un essor important.

Selon ces approches, si les projets d'aide avaient de mauvais résultats, ce ne pouvait donc être que de la faute des pays bénéficiaires de l'aide qui ne mettaient pas correctement en place les réformes conseillées par les donateurs. En aucun cas les politiques d'aide des donateurs étaient sujettes à une remise en cause, au-delà de la question des montants déversés. Nous abordons ici un point important : les champs de l'évaluation se sont toujours élargis en fonction des objets qui faisaient polémiques, sur lesquels les décideurs devaient se protéger de la critique. L'évaluation prit de l'ampleur dans le secteur de l'aide au développement à mesure que les politiques d'aide des donateurs firent l'objet de critiques.

La Déclaration de 1970 sur la 2ème Décennie pour le développement ne réévalua pas les objectifs de la décennie précédente, bien que ceux-ci n'aient pas été atteints[98]. La seule modification consista en une révision de la méthode de mesure des fonds d'APD car, depuis 1969, l'aide au développement était distinguée des transferts de fonds privés. La cible à atteindre passa alors de 1 % du PNB des États donateurs à 0,7 %[99]. L'APD devint alors une

[96] ALMOND, Gabriel. A, COLEMAN, James. S, *The politics of the Developing Areas*, Princeton University Press, 1960, 608p.
[97] BADIE, Bertrand, *L'État importé. Essai sur l'occidentalisation de l'ordre politique*, Fayard, 1987, 334p.
[98] STOKKE, Olav, 2009, *op. cit.* p. 160.
[99] DAC-OECD, 1972, *op. cit.*

action de politique étrangère clairement différenciée des autres secteurs de la diplomatie économique et stratégique des États.

Le bilan médiocre de la 2ème Décennie pour le développement n'entama pas la confiance des donateurs dans leurs politiques de développement, malgré l'essor de mouvements contestataires des pays du Sud au sein des Nations unies. Le mouvement de la *Dependencia*[100] poussa bien les États du Nord à proclamer, en 1974, un « nouvel ordre économique mondial » (NOEI) à l'AGNU. Toutefois, rien ne changea dans l'idéologie de l'aide au développement, qui reposait toujours sur les principes centraux de l'approche de Rostow. En réalité, ce n'était pas vraiment un nouveau régime de la coopération pour le développement que les États du Nord proposèrent à travers le NOEI, mais un nouveau partage des bénéfices de la croissance, sans pour autant changer les moyens de les atteindre[101].

Suite à cette proclamation, la Charte des droits et devoirs économiques des États fut adoptée[102]. Elle légitimait la croissance des budgets d'aide au développement en donnant comme argument principal l'extension de la logique capitaliste de marché mondialisé qui nuisait aux pays du Sud. Cette charte marqua le début de l'approche dite « tiers-mondiste » du développement qui donnait une valeur morale à l'aide[103]. L'aide serait un devoir envers les populations les plus démunies disposant d'un droit à la recevoir[104]. C'est ainsi que le caractère moral et éthique de l'aide fut établi comme une norme internationale, ce qui devait permettre de sanctuariser les budgets d'APD. La Charte réaffirmait implicitement, contrairement à ce que prônaient les Dépendantistes, que c'était au Nord que se trouvait la clé du développement du Sud, puisqu'il portait la responsabilité de la pauvreté des populations du Sud. Cette approche de l'aide ne favorisait pas l'évaluation, dans la mesure où les donateurs n'avaient à justifier, ni l'augmentation des budgets d'APD, ni l'efficacité de leurs actions qui n'était pas remise en cause.

---

[100] Le mouvement de la « Dependancia » porté notamment par Raul Prebish, économiste argentin aux Nations unies, soutenait que les États du Nord tireraient leur croissance économique du sous-développement du Sud, c'est-à-dire de l'exploitation de ses ressources naturelles et de sa main d'œuvre. Cette domination externe serait renforcée par un système de domination interne, par les élites nationales au service des États occidentaux, qui entraînerait des inégalités nationales, des dynamiques d'exclusion, de concentration des revenus et de satisfaction des besoins des classes aisées de la population au détriment des plus démunis par les pouvoirs politiques. Ce ne serait donc pas une intégration au marché capitaliste qui permettrait le développement des pays du Sud comme le théorisaient les pays occidentaux mais, au contraire, la sortie de ce système. Cette orientation intellectuelle développait une vision unifiée et solidaire du Sud de type « centre-périphérie » .Elle connut son apogée lors de la conférence d'Arusha du 5 février 1967, durant laquelle ses défenseurs, principalement des pays d'Amérique latine, présentèrent une vision « autocentrée » du développement, fondée sur l'autonomie des pays et le refus de l'importation de modèles étrangers.

[101] RIST, Gilbert, *Le développement, histoire d'une croyance occidentale*, Presses de Sciences-po, 4ème édition, 2013, 512p, p. 240-244.

[102] Assemblée générale des Nations unies, *résolution 3281*, 12 décembre 1974.

[103] RIST, Gilbert, 2013, *op. cit.* p. 242.

[104] Se référer à l'annexe 2 qui présente, sous forme de tableau, les grandes évolutions de l'histoire politique de l'aide au développement.

Un autre rapport fondateur des années 1970 fut le Rapport sur le développement et la coopération internationale, commandé à la Fondation Dag Hammarskjöld pour la 7ème session extraordinaire de l'AGNU en 1975. Il présentait le développement comme un processus global allant au-delà de la croissance économique et comme une dynamique interne qui naissait de la culture de chaque société[105]. Il marqua une rupture avec la croyance en une formule universelle du développement, ce qui allait dans le sens des mouvements contestataires du Sud. Ce rapport constitua l'acte de naissance de l'approche de l'aide par les « besoins fondamentaux »[106] qui replaçait la population au centre du système de l'aide. Pour assurer un niveau de vie minium à chacun, l'aide devait être dirigée vers les plus vulnérables.

L'approche tiers-mondiste présentait le sous-développement comme un enjeu mondial, touchant les États Sud et, par contagion, les États du Nord et relevant donc de la responsabilité de tous. Cette approche fut relayée par la Banque mondiale qui, en raison de son budget élevé et de la qualité de son expertise économique, se posait comme le chef de file de la réflexion sur l'aide au développement. Robert Mc Namara, alors président de la Banque, développa la notion de « pauvreté absolue », définie par « des conditions d'existence étiolées au point d'empêcher la réalisation du potentiel génétique que chacun de nous porte à sa naissance »[107]. L'aide au développement se justifiait et se légitimait ainsi par des considérations humanistes : aucun individu ne devait rester en état de pauvreté absolue. Elle devait être distributive pour corriger une répartition injuste des ressources[108].

Malgré de nombreuses critiques[109], les concepts de « biens fondamentaux » et de « pauvreté absolue » devinrent de nouveaux référentiels pour évaluer les résultats de l'aide au développement. Ils constituèrent le premier pas vers des évaluations qui dépassaient la sphère purement économique des analyses coûts-bénéfices et qui s'intéressaient davantage aux aspects qualitatifs des politiques d'aide. Toutefois, la création de ces référentiels précéda le

[105] Fondation Dag Hammarskjöld, *Que faire: un autre développement,* préparé à l'occasion de la septième assemblée extraordinaire des Nations unies, 1975.
[106] La nutrition, le logement, la santé, l'éducation et l'emploi.
[107] Mc NAMARA, Robert, *One Hundred Countries, Two Billions People: the dimensions of Development,* Praeger, 1973, 140p.
[108] La notion d'aide distributive est définie en annexe 2.
[109] Les experts en statistiques du développement et en évaluation ont beaucoup critiqué la notion de pauvreté absolue, car les « besoins fondamentaux » ne sont pas des données absolues déterminées par des caractéristiques physiologiques, mais des besoins relatifs construits par un contexte sociétal. Le seuil de survie est difficile à évaluer car il est déterminé de manière relative, par une comparaison des conditions de vie au sein d'une société donnée. Ainsi, les conditions nécessaires à la survie sont souvent inférieures à celles requises pour assurer la vie en société, et la détermination d'un niveau minimal universel de besoins fondamentaux par des acteurs externes aux sociétés concernées est un exercice contestable. Il est toutefois largement reconnu parmi ces détracteurs que ce seuil minimal a le mérite en théorie, d'assurer un « droit » à l'aide aux populations les plus démunies.

développement de capacités techniques et humaines dans les agences d'aide pour en évaluer les progrès.

C'est une vision causale et mécaniste des interventions d'aide qui dominait donc la pensée de l'aide au développement dans les années 1960 et 1970. Les projets étaient mis en œuvre à petite échelle, de manière isolée les uns des autres et portaient sur des thématiques techniques (infrastructure, industrie, etc.) ou économiques. Les résultats étaient donc perçus comme étant facilement maîtrisables et prévisibles.

En conséquence, les rares évaluations qui étaient mises en œuvre avaient pour objectif de confirmer la bonne réalisation des activités d'aide par rapport aux objectifs fixés en amont des projets et surtout aux montants engagés. Très peu de considération était donnée à l'évaluation des processus de mise en œuvre des projets d'aide et de leurs résultats. Nous retrouvons donc l'approche « sommative »[110] de l'évaluation développée aux États-Unis, avec des projets qui restaient des « boîtes noires » pour les évaluateurs. Toutefois, si les pratiques de l'évaluation de l'aide qui émergèrent étaient largement inspirées des modèles développés dans les politiques publiques américaines (approche sommative et vision rationaliste du développement), les experts de la coopération internationale adaptèrent ces approches aux besoins de leurs organisations, notamment en élaborant des outils d'évaluation spécifiques à l'aide au développement.

- **Les premiers instruments de suivi-évaluation de l'aide au développement**

Bien que l'évaluation ait été une pratique marginale dans le secteur de l'aide au développement dans les années 1960, c'est l'époque où les organisations multilatérales commencèrent à mettre en place les premiers mécanismes de suivi[111] des politiques internationales de développement. Les Nations unies jouèrent un rôle précurseur avec l'ECOSOC qui supervisait les activités de l'ensemble des agences onusiennes. Olav Stokke, dans son ouvrage retraçant l'histoire des Nations unies, décrit l'émergence d'une culture de l'évaluation et de la transparence dans cette organisation. Selon lui, elle fut favorisée par l'adoption de la Déclaration de la Décennie pour le développement qui donna des objectifs aux donateurs et aux pays bénéficiaires[112]. Les Nations unies durent, en conséquence, élaborer des

[110] Sont qualifiées de "sommatives" les évaluations qui ont pour objectif non pas un apprentissage par les acteurs impliqués dans la politique publique évaluée, mais l'apport de preuves sur son efficacité. Ce type d'évaluation sert à produire un bilan des résultats de la politique publique menée.
[111] La différence entre suivi et évaluation est expliquée dans le lexique en annexe 7.
[112] STOKKE, Olav, 2000, *op. cit.* pp. 198-203.

instruments pour suivre et évaluer les avancées, globales et par acteur, vers l'atteinte de ces objectifs.

Il faut rappeler qu'à cette époque, les capacités d'analyse économique et sociale dans les pays du Sud étaient très faibles. L'absence de données sur les besoins des populations et sur leur aptitude à soutenir des efforts de développement handicapait nécessairement la planification des programmes de développement et la détermination d'objectifs globaux ou nationaux réalistes, reflétant les besoins et les capacités des pays bénéficiaires. Ainsi, les efforts des Nations unies se concentrèrent, tout d'abord, sur le renforcement des capacités statistiques des pays du Sud, afin de permettre l'établissement d'objectifs adaptés à chaque contexte national et non plus établis de manière théorique, sans prise en compte des spécificités des pays concerné[113]. Ce processus de collecte et de capitalisation de données statistiques permit aux agences onusiennes de prendre conscience de la diversité des conditions entre les pays en développement et d'accumuler des connaissances sur l'état du développement dans le monde.

Dès les années 1960, les Nations unies et ses agences spécialisées commencèrent à produire des rapports périodiques sur les progrès et les tendances du développement économique et social à travers le monde. Au fil des années, les rapports se mirent à comparer systématiquement les progrès réalisés par rapport aux objectifs établis pour la Première Décennie de l'aide. Ce n'est pas tant dans un souci de redevabilité que ces objectifs furent fixés et suivis par les Nations unies, que de montrer que l'ensemble des donateurs était unifié vers l'atteinte d'un ensemble d'objectifs communs. Ces rapports constituèrent également les premières études de comparaison internationale des parcours de développement permettant de faire des analyses macro des facteurs de succès ou d'échecs des projets d'aide.

À partir du début des années 1970, les agences d'aide commencèrent à développer des instruments d'évaluation des projets au niveau microéconomique. C'est aux États-Unis que fut développé, en 1969, l'outil encore aujourd'hui le plus utilisé pour suivre et évaluer les interventions d'aide au développement : le cadre logique, conçu par Leo Rosenberg pour le compte d'USAID[114]. Cet outil de gestion et d'évaluation connut un grand succès à travers le monde, et de nombreuses agences d'aide multilatérales, bilatérales et non gouvernementales se mirent à l'utiliser dès les années 1980. Il devint rapidement un outil obligé du développement. Le cadre logique est défini comme « un ensemble de concepts emboîtés qui peuvent être utilisés

[113] Ce fut pourtant la démarche adoptée 40 ans plus tard par les Nations unies avec la création des Objectifs du Millénaire (voir chapitre 3).
[114] GIOVALUCCHI, François, OLIVIER DE SARDAN, Jean Pierre, « planification, gestion et politique dans l'aide au développement : le cadre logique, outil et miroir des développeurs », *Revue Tiers Monde,* vol 2, n°198, 2009, pp. 383-406.

tous ensemble d'une manière dynamique pour permettre l'élaboration d'un projet bien conçu, objectivement décrit et évaluable. »[115]. En des termes moins technocratiques, le cadre logique sert, sous forme de tableau[116], à décomposer de manière *ex ante* les moyens, réalisations, résultats et impacts des projets d'aide au développement en montrant comment s'établit le lien logique entre chacune de ces étapes. Cet outil permet donc à l'opérateur du projet d'avoir une vision globale de la chaîne de causalité qu'il doit suivre pour atteindre ses objectifs et aide l'évaluateur à comprendre plus facilement où le projet a fonctionné et où il a connu des blocages. Christian Castellanet défend l'idée que « la méthode du cadre logique est centrée sur l'idée qu'une meilleure préparation des projets doit permettre de proposer une planification détaillée des activités avec leur coût et leurs résultats attendus. Le suivi est ensuite grandement facilité, puisqu'il n'y a plus qu'à vérifier que les activités se déroulent conformément aux prévisions, et dans le cas contraire, de prendre des mesures correctives. C'est ce qu'on peut donc appeler des projets clés en main »[117].

Selon François Giovalucchi, ancien économiste à l'AFD, et Jean-Pierre Olivier de Sardan, anthropologue, « le principe de construction (du cadre logique) reste partout identique : il s'agit d'établir, avant une opération de développement, une série de relations logiques entre des objectifs hiérarchisés, les moyens mis en œuvre pour les atteindre, les conditions externes indispensables (hypothèses critiques) et les résultats attendus, tout en identifiant une série d'indicateurs quantitatifs destinés à évaluer les moyens mis en œuvre pour l'opération, ainsi que les réalisations et l'impact de cette opération »[118]. Cet instrument illustre donc la volonté des donateurs de rationaliser leurs stratégies. Selon nous, grâce à cet outil, l'évaluation va trouver sa place dans la gestion de l'aide. Elle va devenir une étape à part entière et non facultative des projets d'aide au développement et donc se systématiser.

Au vu de l'influence importante que l'outil du cadre logique a eu dans le monde de l'aide au développement, il est étonnant que si peu de littérature soit consacrée aux effets qu'il a entraînés sur les pratiques des donateurs. Les documents publiés sur le cadre logique sont principalement des guides méthodologiques de portée technique, internes aux agences d'aide. Très peu d'ouvrages académiques abordent les enjeux politiques qu'a fait émerger cet outil. Or, nous estimons que son utilisation systématique et routinière a joué un rôle significatif dans le développement d'une approche « technique » de l'aide au développement ne prenant pas en

[115] REBIEN, 1996, *op. cit.*
[116] Un exemple de cadre logique est présenté en annexe 1.
[117] CASTELLANET, Christian, « Cycle des projets, cadre logique et efficacité des interventions de développement », *Transverses*, n°13, GRET. 2003, p. 14.
[118] GIOVALUCCHI, OLIVIER DE SARDAN, 2009, *art. cit* p. 388.

compte le contexte politique et social des projets, éléments difficilement intégrables aux cadres logiques des projets. Malgré son utilité en termes de planification, le cadre logique a été un facteur de technocratisation et de « décontextualisation » de l'aide au développement.

L'élaboration de cadres logiques devint rapidement une étape obligatoire de la conception de projets d'aide au développement dans les agences d'aide anglo-américaines. La bonne conception et le suivi du cadre logique étaient considérés comme une condition à la faisabilité des évaluations. La systématisation de cette pratique a été critiquée car, élaborés de manière routinière, les cadres logiques rendraient rigide la gestion des projets, empêchant les opérationnels de s'adapter aux imprévus et de développer des approches innovantes. Bien qu'ils soient toujours omniprésents aujourd'hui dans les « boite à outils » des opérateurs de projets d'aide, ces derniers déclarent souvent les considérer comme un carcan[119]. Toutefois, ils reconnaissent que cet instrument a pour atout de les obliger à réfléchir aux objectifs qu'ils veulent atteindre et aux schémas de causalité qui pourraient faire se réaliser ces objectifs. Le cadre logique permet également un suivi uniformisé et rapproché des projets qui facilite leur évaluation *ex post*. Enfin, l'instrument du cadre logique correspondait bien aux projets d'infrastructure et d'assistance technique répandus dans les années 1960 et 1970, dont la mise en œuvre était linéaire et les résultats relativement prévisibles en amont.

Le cadre logique est une illustration d'un instrument d'évaluation qui s'est diffusé du secteur de l'aide au développement vers les politiques publiques « nationales ». Par exemple, aux États-Unis, le cadre logique a été réutilisé dans le cadre de la méthode du PPBS dans les années 1960. En France, la RCB utilisait le cadre logique afin d'« inscrire la décision budgétaire dans un cercle vertueux 'évaluation *ex ante*, décisions, suivi, évaluation ex post, nouvelles décisions' »[120]. C'est le signe d'une volonté généralisée des décideurs politiques de rationaliser leurs systèmes de gestion de l'administration, préoccupation qui restera un leitmotiv jusqu'à nos jours.

Albert Hirschman fut l'un des premiers théoriciens de l'économie du développement à publier un ouvrage remettant en cause ces méthodes de gestion et d'évaluation des projets d'aide au développement[121]. Selon lui, les analyses coût-bénéfice et les modèles complexes de gestion de l'aide étaient utilisés par les dirigeants des agences de développement pour persuader

[119] Constat issu de témoignages au sein de l'AFD et du DFID ainsi que lors de journées d'étude sur le thème de l'évaluation de l'aide au développement, voir liste des entretiens en annexe 8.
[120] PERRET, Bernard, « De l'échec de la rationalisation des choix budgétaires (RCB) à la loi organique relative aux Lois de Finances », *Revue française d'Administration publique*, n°117, 2006, pp. 31-41.
[121] HIRSCHMAN, Albert. O, *Development Projects Observed*, The Brookings Institution, 1967, ed. 2002, 218p.

les décideurs politiques de la fiabilité de projets qui auraient rationnellement dû être rejetés en raison de difficultés évidentes. Selon lui, ces difficultés auraient été identifiables par des analyses du contexte de mise en œuvre des projets. Nous voyons donc en Albert Hirschman l'un des premiers détracteurs de l'évaluation « scientifique » et des outils tels que le cadre logique, utilisés comme des moyens de planification et d'évaluation de l'aide « en laboratoire », élaborés en dehors de toute influence contextuelle.

Albert Hirschman estimait que les évaluations *ex ante* jargonantes, les cadres logiques et les procédures de suivi servaient à couvrir l'ignorance des experts sur les réelles solutions aux problèmes qu'on leur demandait de résoudre. S'ils savaient réellement comment dépasser les obstacles au développement identifiés dans des contextes nationaux spécifiques, ils proposeraient des solutions bien plus simples, ciblées et adaptées aux contextes des pays[122]. Toutefois, selon lui, si les décideurs connaissaient l'ampleur réelle des difficultés auxquelles devraient se confronter les opérationnels dans la mise en œuvre des interventions d'aide, ils ne s'engageraient jamais dans de tels investissements, car ils sous-estiment leur capacité à résoudre les problèmes. Il est donc préférable que le décideur n'ait pas conscience des risques qu'il prend et qu'il agisse par essais-erreurs lorsqu'il y fait face. Ce n'est que dans un second temps, lorsque les décideurs prennent conscience de leur aptitude à faire face aux enjeux de développement, qu'une évaluation plus mature des projets pourrait être menée, sans qu'il y ait un risque d'abandon. En effet, c'est souvent l'espoir de résultats importants à court terme et en contrepartie d'efforts minimaux qui convaincrait les acteurs politiques de valider des projets plutôt que d'autres[123].

Albert Hirschman fut un des premiers théoriciens à exposer qu'un projet ne devrait pas seulement être évalué *ex post* par rapport à ses réalisations directes ou à son taux de retour sur investissement, mais aussi par rapport aux changements humains et sociaux qu'il engendre et à ses effets indirects positifs et négatifs. Les projets devraient être évalués en fonction de leur capacité à produire des changements en puisant dans les ressources humaines et matérielles présentes dans les pays bénéficiaires. Selon une vision inspirée de l'anthropologie, Albert Hirschman défendait l'idée que l'inefficacité des projets d'aide serait souvent causée par l'ignorance des circonstances locales, des traditions et des structures politico-sociales, ainsi que par le manque de réalisme des objectifs des projets. Toutefois, il soulignait que produire un développement approprié par les pays du Sud demandait une approche de long terme de l'aide

[122] HIRSCHMAN, Albert, O, 1967, *op. cit.* pp. 23-24.
[123] *Ibid.* p. 10.

et que les donateurs acceptent une performance moindre et une prise de risques[124]. Sa vision de l'efficacité de l'aide au développement mettait déjà en avant l'importance de l'appropriation et de responsabilisation des bénéficiaires de l'aide, concepts qui n'ont émergé comme paradigmes centraux des politiques d'aide que dans les années 2000.

Enfin, Albert Hirschman se positionnait contre la quête d'un système unique d'évaluation et de classement des projets. Il trouvait cet exercice futile, dans la mesure où les projets renfermeraient des dimensions trop multiples pour être agrégées dans un indice unique, bien que cette simplification didactique puisse aider à la décision. Nous retrouverons, dans toute l'histoire de l'évaluation, ce dilemme entre une simplification des critères d'évaluation pour aider à la prise de décision et une analyse complexe des projets, passant par l'élaboration de matrices multifactorielles et la collecte de données qualitatives pour améliorer la compréhension des enjeux du développement[125].

Ces apports d'Albert Hirschman à la théorie de la gestion et de l'évaluation des projets d'aide ont été ignorés pendant de nombreuses années et n'eurent que peu d'impact sur les pratiques des donateurs à cette époque. Ce ne fut qu'à la fin des années 1980 que les penseurs du développement redécouvrirent ces travaux et y trouvèrent une valeur analytique utile à la réforme de la gestion et de l'évaluation des activités d'aide, après que l'approche de l'aide néolibérale du Consensus de Washington ait échoué.

B. Le tournant des années 1980 : l'idéologie néolibérale de l'évaluation

Nous allons désormais aborder une deuxième phase de l'histoire politique de l'évaluation, de 1980 à 1991, période qui marqua un tournant néolibéral pour l'aide au développement et l'évaluation. L'évaluation va devenir un outil de rationalisation des pratiques d'aide et de conditionnalité[126] envers les pays bénéficiaires dans un contexte idéologique où l'aide n'est plus considérée comme un devoir des donateurs, mais comme un privilège que les pays bénéficiaires doivent mériter. Les pays bénéficiaires sont considérés comme responsables de leur développement et doivent s'engager à mettre en place des réformes dont la réussite sera évaluée par les donateurs et conditionnera les montants d'aide qu'ils recevront.

[124] HIRSCHMAN, Albert. O, 1967, *op. cit.* pp. 130-135.
[125] *Ibid.* pp. 175-190.
[126] Pour une définition du concept de conditionnalité appliqué à l'aide au développement, voir lexique en annexe 7.

- **Remises en causes de l'efficacité de l'aide et conditionnalités**

Jusqu'à la fin des années 1970, les fonds d'aide augmentèrent continuellement d'année en année. Lorsque les États occidentaux effectuaient des coupes budgétaires, les politiques d'aide étaient « sanctuarisées » du fait de leur caractère humaniste et éthique. Toutefois, cette situation évolua avec la crise économique, touchant de manière généralisée les pays occidentaux à la suite du choc pétrolier de 1979, qui entraîna une crise monétaire aux États-Unis en 1980 et une montée des idéologies néolibérales[127] à travers le monde, notamment dans les pays anglo-américains.

Ces pays allouèrent alors leurs fonds d'aide de manière plus sélective en fonction des perspectives de croissance économique des États. L'aide n'était plus considérée comme un droit universel et devait être distribuée plus rationnellement selon une logique méritocratique. Basil Cracknell, évaluateur à l'ODA dans les années 1980, explique que « pour la première fois, les agences d'aide durent justifier leurs choix d'allocation auprès de leurs tutelles financières et prouver leur efficacité »[128]. Nous identifions ici un changement majeur par rapport à la décennie précédente, qui modifia profondément le rapport des acteurs de l'aide à l'évaluation. Les décideurs des organismes d'aide eurent désormais besoin de l'évaluation, y voyant un outil leur permettant de justifier leur choix d'allocation de l'aide et de prouver la performance de leurs activités.

Les volumes d'aide des États-Unis, du Royaume-Uni, du Danemark et des Pays-Bas, ces deux derniers pays étant imprégnés par une culture anglo-saxonne des politiques publiques, furent en baisse durant la première moitié des années 1980. Au Royaume-Uni, les volumes d'aide passèrent de 2 157 millions de dollars en 1979 à 1 429 millions de dollars en 1984, au Danemark de 481 millions de dollars en 1980 à 394 millions en 1983 et aux Pays-Bas de 1 630 millions de dollars en 1980 à 1 135 millions en 1985. Aux États-Unis, l'aide passa de 5 664 millions de dollars en 1978 à 4 684 millions de dollars l'année suivante[129]. Cette baisse des crédits pouvait s'expliquer par les phénomènes croisés de contrainte budgétaire et de politiques de rationalisation des allocations d'aide, inspirées par l'idéologie montante du néolibéralisme. En France, par contre, pays où la perception humaniste de l'aide était plus

[127] Selon Michel Foucault (*la Naissance de la biopolitique, cours au Collège de France*, publié en 2004), le néolibéralisme est présenté comme une technique de gouvernement, une politique économique et sociale étendant l'emprise des mécanismes du marché à l'ensemble de la vie. Dans les années 1980, Margaret Thatcher et Ronald Reagan s'inspirèrent des théories néolibérales pour réformer leurs politiques publiques vers une réduction de l'emprise du secteur public sur le secteur économique et une réduction de la bureaucratie administrative pour une meilleure efficacité et une réduction des coûts de la fonction publique.
[128] CRACKNELL, 2000, *op. cit*, p.44.
[129] Données tirées des statistiques du CAD-OCDE: http://stats.oecd.org/qwids/, consulté le 7 novembre 2014.

ancrée et où les théories néolibérales étaient moins populaires, le budget consacré à l'APD resta stable et la pratique de l'évaluation peu développée. À ce contexte, s'ajoutait une crise de la dette dans un grand nombre de pays du Sud, qui incita les bailleurs de fonds à conditionner davantage leurs prêts et donc à diminuer leurs montants d'aide.

Des chercheurs entreprirent alors de développer des méthodes scientifiques d'évaluation de l'efficacité macroéconomique de l'aide. La plupart d'entre eux, le plus souvent anglo-américains, travaillaient parallèlement à leur carrière académique dans des organismes de développement. Ce double statut leur donna l'opportunité de susciter un intérêt sur la recherche en évaluation, aussi bien dans leurs universités, qu'auprès des « praticiens » de l'aide au développement. Le plus éminent de ces chercheurs fut Paul Mosley, universitaire de carrière qui fut conseiller économique à l'ODA à la fin des années 1970, puis consultant dans des organisations multilatérales. Il fonda, au début des années 1980, un groupe de pression baptisé *Independent Group on British Aid* qui proposait aux décideurs britanniques des réformes du système de l'aide au développement.

Ces chercheurs utilisaient l'évaluation comme un instrument de critique et de débat éclairé sur l'efficacité de l'aide. Leur objectif consistait à rendre l'aide au développement plus transparente et à mettre ses dysfonctionnements en lumière. L'évaluation, jusqu'alors confinée à l'expertise opérationnelle des organismes d'aide, entra dans la sphère de la recherche. Nous avons observé que cette évolution conduisit à « intellectualiser » la fonction d'évaluation. Le fait que les chercheurs se soient saisis de la question de l'évaluation contribua à la crédibiliser en tant qu'objet académique de réflexion sur le développement.

Paul Mosley publia, notamment, deux rapports d'évaluation de l'efficacité macroéconomique de l'aide qui se basaient sur un grand nombre d'évaluations produites par les agences d'aide à travers le monde. Sa première étude[130], en 1981, portait sur l'évaluation de l'efficacité de l'aide quant à la réduction de la pauvreté. En 1986[131], il publia une étude qui tentait d'évaluer les résultats de l'aide en général[132]. Dans ces travaux, il démontra l'existence

[130] MOSLEY, Paul, *Aid as obstacle: Twenty questions about our foreign aid and the hungry,* Institute for Food and Development Policy, 1980, 192p.

[131] MOSLEY, Paul, "The Political Economy of Foreign Aid: A Model of the Market for a Public Good," *Economic Development and Cultural Change*, vol. 33, n°2, January 1985, pp. 373-93.

[132] Ces deux rapports utilisaient la méthode statistique de la régression linéaire. Elle permet d'évaluer que la hausse d'une variable (par exemple, l'aide au développement) est associée à un effet plus ou moins important à la hausse ou à la baisse d'une autre variable (par exemple, la croissance économique). Appliquée à plusieurs dizaines de pays en développement, cette méthode permet d'observer une corrélation entre l'aide au développement, ou des interventions plus spécifiques, et une augmentation de la croissance économique. Cette mesure se fait par exemple par un calcul de type : le rapport de la croissance du produit intérieur brut (PIB) par habitant et l'aide au développement rapportée au PIB, observé pour N pays durant une certaine période. Plus le ratio est proche de 1 plus la corrélation est forte.

d'un écart d'efficacité entre l'échelle « micro » des projets, où l'aide apparaissait efficace, et l'échelle globale où elle apparaissait inefficace. Ces travaux furent à l'origine de l'expression « micro-macro gap » qui désigne, aujourd'hui encore, l'incapacité des activités d'aide à généraliser leurs bons résultats à une échelle plus large que celle du projet. Ces deux rapports furent les premières études qui remirent en cause de manière « scientifique » l'efficacité de l'aide. Peu de temps après, un autre travail de recherche marquant fut publié par Roger Riddell[133], chercheur à *l'Overseas Development Institute* (ODI) qui mettait en miroir les arguments « pour » et « contre » l'aide au développement. Son ouvrage fut perçu comme ayant une forte valeur ajoutée, car il fondait son analyse sur un grand nombre d'évaluations quantitatives des résultats de projets d'aide menées par divers organismes d'aide.

Nous considérons que les polémiques sur l'efficacité de l'aide provoquées par ces travaux furent le premier signe avant-coureur que cet enjeu allait devenir majeur dans les décennies suivantes. Ces travaux amorcèrent une réflexion sur l'efficacité macroéconomique de l'aide, alors que les organismes d'aide (excepté la Banque mondiale) se cantonnaient jusqu'alors à l'évaluation de son efficacité microéconomique à l'échelle des projets, en raison de leur expertise limitée. L'efficacité microéconomique de l'aide étant plus facile à atteindre et à prouver et le projet étant l'unité de mise en œuvre de l'aide la plus courante, les agences d'aide n'avaient pas d'intérêt à s'investir dans des évaluations de son efficacité macroéconomique en l'absence d'une demande externe.

Ces travaux de recherche, très critiques, marquèrent le début d'une période de remise en cause de l'efficacité de l'aide qui a perduré jusqu'à aujourd'hui. Dès lors, les travaux cherchant à critiquer ou à réhabiliter l'efficacité de l'aide se succédèrent sans cesse, dénonçant, pour justifier leur valeur-ajoutée, le manque de rigueur des méthodes utilisées par leurs prédécesseurs[134]. Ces travaux ont tendance à traiter de manière élusive les problématiques de fond sur la qualité de l'aide pour se restreindre à des débats technocratiques entre différents courants méthodologiques. Les évaluations réalisées par les chercheurs externes aux agences d'aide s'éloignèrent ainsi de leur vocation d'analyse politique pour se limiter à une confrontation technique des différentes méthodologies d'évaluation de l'efficacité de l'aide.

Face à cette remise en cause de l'efficacité de l'aide, les donateurs firent évoluer les normes et les valeurs du régime de l'aide. Les années 1980 marquèrent une rupture avec le principe « tiers-mondiste », pour laisser place à un régime « néolibéral » de l'aide, où l'efficacité

[133] RIDDELL, Roger. C, *Foreign Aid Reconsidered*, the John Hopkins University Press, 1987, 320p.
[134] Ces principaux travaux sont présentés en annexe 3.

des interventions était principalement mesurée en termes de rentabilité financière pour les bailleurs de fonds et de croissance économique pour les pays bénéficiaires[135]. Selon cette approche néolibérale, les pays du Sud portaient la responsabilité de leur situation de sous-développement[136]. Cette dernière n'était plus considérée comme la résultante de conditions géographiques et historiques défavorables, mais de mauvaises pratiques économiques.

Cette approche du sous-développement justifiait que l'aide soit allouée à condition que les gouvernements bénéficiaires acceptent de mettre en place des réformes économiques imposées par les donateurs : les « prêts d'ajustement structurels »[137]. L'aide devenait un privilège méritocratique pour récompenser les États qui faisaient preuve de « bonne volonté » et d'efficacité à réformer leurs politiques nationales.

Avec l'apparition d'une aide plus conditionnelle naquit le besoin de construire des indicateurs pour sélectionner de manière objective, homogène et transparente les pays qui pouvaient bénéficier de plus ou moins d'aide. En effet, les donateurs cherchèrent à se dédouaner de leur prise de décision politique en se réfugiant derrière des calculs scientifiques plus difficilement critiquables, car perçus comme rationnels et neutres. En conséquence, les dirigeants des organismes d'aide portèrent plus d'intérêt à l'évaluation, car elle devint un moyen de justification rationnelle de la répartition géographique et sectorielle de leur allocation d'aide.

D'après des entretiens menés avec des praticiens et des évaluateurs de l'aide au développement dans les années 1980[138], ce n'était pas la politique des bailleurs de fonds qui était remise en question, mais la manière dont les pays du Sud se « donnaient les moyens » de se développer grâce à l'aide reçue. L'aide au développement n'était qu'un soutien temporaire aux États pour qu'ils puissent se donner eux-mêmes les moyens de se développer. Une autre difficulté serait le manque d'investissement des gouvernements et des populations dans les projets et programmes apportés par les donateurs[139]. Les organismes d'aide et la société civile ne remettaient donc pas en cause l'efficacité de l'aide, mais celle des États du Sud.

Ces derniers s'opposèrent à l'approche néolibérale de l'aide au développement. En 1987, la Commission Sud, créée à l'initiative du Premier ministre malaisien Mohamad et du Président

[135] STOKKE, Olav, 2009, *op. cit.* pp. 315-320.
[136] Terme emprunté à Jean-David Naudet dans *Les OMD et l'aide de cinquième génération : Analyse de l'évolution des fondements éthiques de l'aide au développement, Agence française de développement*, document de travail, 2005.
[137] Ensemble de dispositions dont certaines agissent sur la conjoncture et d'autres sont structurelles, qui résultent d'une négociation entre un pays demandeur d'aide et le FMI ou la Banque mondiale pour modifier le fonctionnement économique du pays, les institutions internationales conditionnant leur aide à la mise en place de réformes qu'elles considèrent pérennes.
[138] Entretien avec un ancien responsable du département de l'évaluation de l'AFD, 21 décembre 2010 ; entretien avec un évaluateur retraité du ministère de la coopération français, désormais consultant auprès d'organisations internationales, 21 mars 2013 ; entretien avec un ancien évaluateur au ministère de la coopération français, 4 avril 2013.
[139] REBIEN, 1996, *op. cit.*

tanzanien Nyerere, réaffirma le principe d'autosuffisance aux Nations unies[140]. Le développement était présenté comme « un processus par lequel les pays et les peuples comptent sur eux-mêmes et décident eux-mêmes des buts à atteindre. Les gouvernements doivent s'interroger sur ce qu'il faut produire, par quels moyens et à quel coût social et écologique »[141]. Les États du Sud devaient s'approprier les programmes d'aide au développement qui nécessitaient des réformes structurelles importantes des sociétés impossibles à imposer de l'extérieur. Par ailleurs, la Commission réaffirma l'idée d'un développement centré sur les individus. Ce mouvement, pour des raisons différentes des néolibéraux, prônait également l'idée d'une meilleure implication des pays du Sud dans l'aide au développement.

Cette émergence de la question de la participation amena les experts du développement à élaborer des méthodes plus participatives de gestion et d'évaluation de l'aide.

- **Rendre l'évaluation plus participative, des tentatives avortées**

Robert Chambers, chercheur à l'ODI, fut l'un des premiers à théoriser l'approche participative de la gestion de l'aide. De 1983 à 1988, il mena une série d'études soulignant l'importance de l'appropriation et de la participation des bénéficiaires dans les projets d'aide au développement[142]. Selon lui, l'inefficacité des projets d'aide serait due à l'insuffisante participation des bénéficiaires dans leur planification et leur mise en œuvre. Toutefois, il ne considérait pas, à la manière les néolibéraux, que la plus grande participation des pays du Sud se justifiait par leur responsabilité dans leur état de sous-développement, mais plutôt par l'inefficacité des pratiques visant à imposer à des populations les modes de vie qu'elles doivent adopter. Imposition qui serait inefficace car elle entraînerait la résistance des populations ciblées. La méthode qu'il développa et qui servit de référence durant plusieurs décennies était intitulée *rapid rural appraisal* (RRA)[143]. Elle avait pour objectif de rendre les populations locales capables d'analyser elles-mêmes leurs besoins, ainsi que de planifier et de mettre en œuvre des actions de développement correspondantes.

Pour Robert Chambers et ses collaborateurs, comprendre la culture des bénéficiaires de l'aide permettrait de choisir des projets mieux adaptés à leurs modes de vie et à leurs valeurs, qui ne seraient pas forcément les plus optimaux d'un point de vue de rationalité économique,

[140] STOKKE, 2000, *op. cit.* pp. 180-192.
[141] *Ibid*.
[142] CHAMBERS, Robert, *Rural Development: Putting the Last First*, Longmans, 1983, 256p ; CHAMBERS, Robert, *Farmers First: A Practical Paradigm for Third World Agriculture*, Institute of Development Studies, Brighton, Sussex, 1988.
[143] CHAMBERS, Robert, 1983, *op. cit.*

mais qui auraient le mérite de pouvoir être appropriés par les populations. La méthode du RRA reposait sur la collecte du savoir local grâce à des techniques innovantes de collecte de données (groupes de réflexion avec les populations locales, entretiens individuels, cartographie sociologique, participation et information par le biais de l'iconographie) pour assurer la participation de tous. Elle avait pour objectif « un renversement de l'apprentissage », des bénéficiaires vers les donateurs. Les fondements de la méthode RRA furent repris par les organisations internationales, notamment le CAD-OCDE qui, en 1986, conclut qu'« une plus grande importance devrait être donnée à s'assurer de l'engagement des agences bénéficiaires exécutantes et de la motivation des groupes ciblés locaux à travers leur implication active dans la sélection, l'élaboration et la mise en œuvre des projets »[144].

Du côté de la discipline des politiques publiques, des chercheurs développèrent des approches participatives de la gestion organisationnelle, qui visaient à replacer les « acteurs » au centre de cette gestion[145]. L'universitaire et évaluatrice Carol Weiss travailla à la transposition de cette approche dans le domaine de l'évaluation des politiques publiques, fondant la théorie de « l'évaluation participative », en opposition aux approches « conventionnelles » de l'évaluation qui n'incluaient que les gestionnaires de projets, et non leurs bénéficiaires[146]. Carol Weiss critiquait le manque d'indépendance des évaluations, leur inutilité pour les acteurs des politiques publiques, leur champ trop étroit et leur faible pouvoir de transformation sur les politiques publiques[147]. Les évaluateurs s'intéressaient de plus en plus à l'utilisation qui serait faite de leurs travaux[148].

Selon Carol Weiss, les recommandations des évaluations ne seraient légitimes que si elles sont utiles à la prise de décision. Si l'on suit cette réflexion, l'objectif de l'évaluation serait plutôt de produire une information utile, qu'une information vraie, comme c'était le cas durant les décennies précédentes. Les évaluations devraient moins se consacrer à déterminer les enjeux de causalité (une action A mène à un résultat B) qu'à comprendre ce qui a bien et mal fonctionné dans le processus de mise en œuvre des projets. Cette théorie penchait donc en faveur d'une utilisation formative[149] des évaluations et incitait les évaluateurs à ouvrir la « boîte noire » des programmes de politiques publiques.

---

[144] DAC-OECD, *Rapport de bilan sur l'évaluation dans les agences d'aide,* Paris, 1986.
[145] FREEMAN, Edward. R, *Strategic Management: a Stakeholder Approach,* Cambridge University Press, 2010, première publication en 1984, 292p.
[146] WEISS, 1983, *art. cit.*
[147] *Ibid.*
[148] WEISS, Carol, 1983, *art. cit.* ; PATTON, 1978, *op. cit.*
[149] Les évaluations "formatives", opposées aux évaluations "sommatives", sont menées dans un objectif d'apprentissage par les acteurs de la politique publique évaluée. Elles vont au-delà de la démonstration de résultats et tentent de comprendre

Nous observons, durant la deuxième moitié des années 1980, la diffusion à travers le monde occidental, d'une approche « pluraliste » et « démocratique » de l'évaluation des politiques publiques, approche jusqu'alors cantonnée aux États-Unis, où elle était minoritaire. Selon Egon Guba et Yvonna Lincoln[150], nous entrons alors dans la quatrième génération de l'évaluation[151], durant laquelle les opérateurs et les bénéficiaires des politiques publiques évaluées devaient participer aux évaluations.

Lorsque nous étudions la littérature sur l'évaluation des politiques publiques de l'époque[152], nous remarquons la naissance d'un consensus sur l'idée que les évaluations devaient émettre des jugements de valeur pour guider les acteurs des projets et des programmes évalués. En France par exemple, le rapport Viveret de 1989 consacra cette approche de l'évaluation. L'auteur fit un plaidoyer contre les « approches techniciennes et positivistes » de l'évaluation qui prétendaient « dire le vrai » sur une politique. Il défendit l'idée que ce ne sont pas les normes et la technocratie qui devraient orienter les politiques publiques, mais les acteurs politiques. Il définissait l'évaluation comme « une conduite destinée à porter un jugement sur la valeur d'une politique »[153]. L'évaluation devrait systématiquement permettre un débat démocratique et donc être transparente dans ses méthodes, ses informations et ses critères de jugement. Il mit, enfin, l'accent sur le fait que l'évaluation devrait être indépendante et mêler des méthodologies qualitatives et quantitatives. La France, qui s'intéressa tardivement à l'évaluation, se mit donc à développer une approche « participative », alternative à la vision néolibérale prédominante aux États-Unis. L'évaluation participative a été conçue avec l'objectif d'aider l'individu à s'auto-assister. Il s'agit d'un processus démocratique, dans lequel les individus renforcent leurs capacités à se forger une opinion et à peser sur les politiques publiques qui les touchent grâce à l'aide d'un expert extérieur faisant principalement office d'encadreur et d'animateur[154].

Toutefois, si la production intellectuelle sur l'évaluation participative a été importante dans les années 1980, les évaluations restaient « très largement technocratiques et maîtrisées

---

pourquoi les résultats ont été atteints ou non et de déterminer, en collaboration avec les parties prenantes de la politique publique, des voies d'amélioration envisageables.

[150] GUBA, Egon. G, LINCOLN, Yvonna. S, "The countenances of fourth generation evaluation: description, judgment and negotiation" in PALUMBO, Dennis (dir.), *The politics of program evaluation*, SAGE Publication, 1987, 309p.

[151] 1re génération dans les années 1960 : l'évaluation « rationnelle » porteuse de vérité pour des acteurs rationnels, 2ème génération au début des années 1970 : remise en cause de l'approche rationnelle, les évaluations doivent guider la prise de décision et produire des jugements de valeur, 3ème génération à la fin des années 1980 : les deux approches précédentes sont justifiables selon les contextes et la nature des politiques publiques évaluées, 4ème génération à la fin des années 1980 et début des années 1990 : l'évaluation doit être participative en intégrant toutes les parties prenantes des politiques publiques évaluées.

[152] Voir bibliographie.

[153] VIVERET, Patrick, *L'évaluation des politiques et des actions publiques*, La Documentation française, 1989.

[154] Pour une définition plus approfondie de l'évaluation participative, voir lexique en annexe 7.

par les gestionnaires des programmes »[155]. Selon un évaluateur retraité du ministère de la coopération, l'émergence de l'évaluation participative a été tardive et chaotique : « *la participation des bénéficiaires est longue, difficile et progressive. Elle n'avait aucun sens tant que l'on ne leur donnait pas la capacité d'influencer la conduite de l'évaluation par le récit de leur expérience* »[156].

Au Royaume-Uni, la pratique de l'évaluation, dans un cadre de restrictions budgétaires, se concentrait sur l'analyse coût-bénéfice des politiques à des fins de rationalisation financière. Une mesure phare, inscrivant l'évaluation dans le contrôle de l'utilisation des fonds publics, fut la création par Margaret Thatcher, en 1979, de *l'Efficiency Strategy Unit* qui avait pour fonction « d'identifier et d'éliminer les gaspillages des administrations ». L'objectif affiché était de « responsabiliser » les fonctionnaires[157]. Afin d'améliorer la performance de l'État, en 1981, le programme *Next Steps* mit en place une « agencification » de l'État, c'est-à-dire un transfert des prérogatives de l'administration à des « agences spécialisées », dotées d'autonomie managériale et contrôlées principalement par des contrats d'objectifs[158]. La création de ces agences devait tester la possibilité d'un transfert de certaines activités publiques vers le secteur privé. Dans le cadre de ce programme, l'évaluation était utilisée comme un outil de comparaison de l'efficience des organismes les uns par rapport aux autres, ce qui mènerait supposément à une mise en concurrence et à une amélioration des services. Cette approche de l'évaluation s'éloignait donc radicalement d'un objectif d'ouverture démocratique et d'amélioration des politiques publiques.

Deux mouvements parallèles sont donc observables dans les années 1980, d'une part la remise en cause des pratiques *top-down* de l'aide au développement et, d'autre part la théorisation d'approches participatives de l'évaluation des politiques publiques. Ils avaient en commun de vouloir remettre les « acteurs », les « bénéficiaires », au centre de l'analyse et de la prise de décision. Cette convergence montre une montée en puissance, à cette époque, de courants de pensée sociologique qui soulignaient l'importance des logiques d'acteurs dans les systèmes organisationnels et politiques[159]. Ces deux mouvements contribuèrent, à la fin des

---

[155] LACOUETTE-FOUGERE, Clément, LASCOUMES, Pierre, *Les scènes multiples de l'évaluation : les problèmes récurrents de son institutionnalisation,* LIEPP, mai 2013. p. 28.

[156] Entretien avec un ancien évaluateur au ministère de la Coopération français, consultant auprès d'organisations internationales, 4 avril 2013.

[157] JACOB, Steve, VARONE, Steve, « Institutionnalisation de l'évaluation et nouvelle gestion publique : un état des lieux comparatif », *Revue internationale de politique comparée,* vol 11, février 2004, pp. 271-292.

[158] BEDECARRATS, Florent, *La microfinance entre utilité sociale et performances financières : le rôle des normes dans la gouvernance d'un secteur mondialisé,* Thèse dirigée par Madame le professeur Josépha Laroche, Université Paris I Panthéon Sorbonne, École doctorale de science politique, 22 juin 2012, 472p.

[159] Par exemple, CROZIER, Michel, FRIEDBERG, Erhard, *L'acteur et le système : les contraintes de l'action collective*, Édition du Seuil, 1981, 500p.

années 1980, aux premières remises en cause directes des méthodes dites scientifiques, ou expérimentales, d'évaluation de l'aide. En 1986, le CAD-OCDE écrivit dans son Rapport de bilan sur l'expérience de la pratique de l'évaluation dans les agence d'aide : « clairement, le modèle classique du format de recherche scientifique expérimental ou quasi expérimental est très rarement, voire jamais, applicable à l'évaluation des projets de développement »[160]. Il y eut alors une prise de conscience parmi les experts en évaluation sur le fait que l'aide au développement ne se prêtait pas à l'idéal type d'un système d'évaluation scientifique et universel qui apporterait des conclusions fiables et non sujettes à la controverse. En effet, dans le domaine du développement, les chaînes de causalité des activités menées sont difficilement maîtrisables, échappant partiellement au contrôle des donateurs. Cette remise en cause de la pertinence des évaluations expérimentales fut vite oubliée et, comme nous le verrons, elles connurent un regain de popularité dans les années 2000, montrant, encore une fois, la faible mémoire institutionnelle dont font preuve les institutions d'aide au développement.

Dans le secteur de l'aide au développement, le plaidoyer des experts en évaluation en faveur des méthodes participatives monta en puissance à la fin des années 1980, lorsque l'efficacité des principes tutélaires du Consensus de Washington commencèrent à être largement remis en cause, menant à qualifier les années 1980 de « décennie perdue »[161] pour le développement.

Les évaluations participatives furent présentées comme un outil permettant de remédier à l'échec, de plus en plus reconnu, de l'aide au développement. Elles permettaient également de répondre à la critique montante de la faible implication des États du Sud dans l'aide au développement. Les évaluateurs des agences d'aide élaborèrent des guides méthodologiques montrant comment appliquer les méthodologies participatives d'évaluation des politiques publiques aux pays du Sud[162]. Les donateurs crurent qu'en faisant participer les autorités locales et les populations à l'évaluation des projets, il serait possible de corriger le fait de ne pas les avoir impliquées dans leur mise en œuvre. Il apparut néanmoins rapidement que faire débuter la participation des populations au moment de l'évaluation ne permettait pas une réelle implication et appropriation des activités d'aide. Une approche plus idéaliste se développa également parmi les évaluateurs, selon laquelle la participation était une fin en soi[163].

[160] OCDE, 1986, *op. cit.*
[161] STOKKE, Olav, 2000, *op. cit.* pp. 315-340.
[162] Voir, par exemple, KUMAR, Krishna, *Conducting key informant interview in developing countries*, USAID, 1989.
[163] RIEBEN, 1996, *op. cit.*

Toutefois, « l'évaluation participative » est un terme conceptuellement faible[164]. Il n'y a toujours pas aujourd'hui de définition harmonisée et reconnue de ce concept parmi les acteurs de l'aide[165]. La participation peut tout désigner : la participation des bénéficiaires directs de l'aide, de la population en général, ou encore des autorités locales. Face à ce flou conceptuel, plusieurs questions apparaissent : Qui devrait participer ? Qui définit les acteurs devant participer ? Quels niveaux d'implication active et d'initiative sont requis pour qu'il s'agisse de participation ? Comment le problème de la variation des ressources et du pouvoir de négociation des parties prenantes peut être géré ? La réponse à ces questions est encore en débat aujourd'hui[166].

Tout comme dans les autres secteurs de politiques publiques, malgré de nombreuses réflexions, la pratique de l'évaluation participative restait quasi inexistante dans l'aide au développement. Ce décalage révélait une tension entre la manière dont de plus en plus d'évaluateurs percevaient leur « mission » (rendre l'aide au développement plus égalitaire, participer à l'autonomisation des populations du Sud, permettre une accumulation de connaissances sur le développement), et ce qu'attendaient d'eux les dirigeants des organismes d'aide au développement (mesurer l'efficacité des projets d'aide et leur donner des arguments objectif pour mieux conditionner les fonds d'aide).

Nous notons seulement deux expérimentations d'évaluations participatives à cette époque, réalisées par la Banque mondiale et la FAO. La Banque mondiale mit en œuvre, dès le début des années 1980, une évaluation participative de grande ampleur du projet d'irrigation Muda en Malaisie, afin de répondre à la question : « qui sont les bénéficiaires et à quel point les bénéfices du projet se diffusent en amont et en aval ? » Cette démarche marqua l'histoire de l'évaluation de l'aide, car ce fut la première à s'intéresser aux impacts des projets d'aide au développement, c'est-à-dire à leurs effets de long terme et indirects[167]. L'évaluation prouva que les travailleurs sans terre bénéficièrent au moins autant des effets du projet d'aide que les agriculteurs spécifiquement ciblés par le projet, en raison de la hausse de la productivité, des revenus et de la consommation de ces derniers, ce qui eut des effets positifs sur leurs niveaux de vie (hausse du commerce, meilleure alimentation, etc.). Cette étude montra donc

[164] WEISS, Carol, 1983, *art. cit. ;* BRUNNER et GUZMAN, 1989, *art. cit.* ; REBIEN, 1996, *op. cit.*
[165] Constat apporté par les conversations informelles menées avec des évaluateurs travaillant dans le secteur de l'aide au développement lors de séminaires ou de journées d'études sur l'évaluation. Voir liste des évènements observés en annexe 9.
[166] FORSS, Kim, *Participatory Evaluation: Questions and Issues*, A Report Commissioned by the Central Evaluation Office, Occasional Paper no. 1. New York, UNDP, octobre 1989.
[167] *Ibid.*

l'importance de s'intéresser aux bénéficiaires indirects des projets pour avoir une vision juste de leurs impacts[168].

La FAO réalisa également, en 1989, des évaluations participatives sur des projets de développement forestier[169]. Ces évaluations avaient pour objectif : i) d'assurer une meilleure implication des catégories les plus pauvres de la communauté dans les activités d'évaluation ; ii) de transformer les rapports sociaux en « augmentant la conscience d'elles-mêmes des populations ». Toutefois, ces premières expérimentations d'évaluations participatives connurent plusieurs difficultés : i) poser des questions évaluatives aux populations sans biaiser les réponses apparut difficile. Les bénéficiaires interrogés avaient tendance à donner les réponses qu'ils estimaient bonnes aux yeux des bailleurs ou des autorités locales ; ii) les dirigeants locaux étaient réticents quant à l'organisation de ce type de consultations avec des populations vulnérables. Cette pratique étant peu habituelle à l'époque, elle nécessitait une sensibilisation préalable des autorités locales, surtout dans des pays où l'expression de la société civile était étroitement contrôlée[170].

Hormis ces premières expérimentations, les évaluations participatives restaient, dans les années 1980, un élément purement rhétorique[171]. Ce n'est qu'à la fin des années 1990 que les évaluateurs, et plus largement les donateurs dans leur ensemble, se saisirent pleinement de l'enjeu de la participation et développèrent des moyens pour la mettre en œuvre. Nous faisons l'hypothèse que s'il a fallu attendre presque vingt ans pour que la participation et la responsabilité mutuelle du développement deviennent des enjeux majeurs de l'aide au développement, c'est parce que ces enjeux étaient perçus comme étant de second plan et trop complexes à traiter par la majorité des praticiens de l'aide. De plus, l'approche néolibérale portée par la Banque mondiale et les organismes d'aide anglo-américains était en contradiction avec une plus grande autonomie des États du Sud dans leurs choix de développement. Le rapport de force était trop déséquilibré en défaveur des porteurs d'une approche sociale et participative de l'aide au développement, ce qui fut mal vécu par les évaluateurs qui pensaient

[168] À cette époque, l'évaluation « d'impact » était une démarche qualitative (qui pouvait être complétée par des éléments d'analyse quantitatifs) qui visait à interroger les bénéficiaires sur les effets qu'avaient eu les projets sur leur vie. Au contraire aujourd'hui, le terme « évaluation d'impact » est souvent confondu avec celui d'« évaluation expérimentale ». Il y a donc un « resserrement » de ce que l'on considère comme une évaluation d'impact, comme s'il ne pouvait être validement évalué que par des méthodes quantitatives complexes dans lesquelles la « parole » des bénéficiaires est sous-estimée.
[169] FAO, *Participatory Assessment, Monitoring and Evaluation: A Field Manual*, Community Forestry Unit, FAO, 1989.
[170] CRACKNELL, 2000, *op. cit.* pp. 335-337.
[171] Entretien avec le directeur d'un cabinet d'évaluation des politiques publiques basé en France, 12 avril 2013.

que la participation entre donateurs et bénéficiaires étaient la clé d'une meilleure efficacité de l'aide[172].

Cette exclusion des gouvernements et des bénéficiaires des évaluations montre qu'elles étaient centrées sur les besoins des donateurs et destinées à n'être utilisées que par eux. Les questions examinées ciblaient les intérêts et les préoccupations des donateurs, ou dans le meilleur des cas, les préoccupations des pays partenaires telles que les percevaient les donateurs. L'évaluation n'était donc pas vue par les donateurs comme ayant vocation à augmenter le dialogue entre donateurs et bénéficiaires ou à permettre un apprentissage mutuel.

Les évaluateurs s'opposaient à l'idée selon laquelle l'exclusion des pays bénéficiaires des processus d'évaluation était seulement le fait de leur « égoïsme ». Ils soulignaient que les gouvernements partenaires n'étaient pas très demandeurs de participer aux évaluations[173]. Les évaluateurs qui travaillaient pour des organismes d'aide ouverts à l'idée d'intégrer les bénéficiaires aux évaluations témoignaient qu'ils devaient créer la demande et mettre en place des activités de renforcement de capacités. Ainsi, des évaluateurs de Suède et du Danemark témoignèrent que les résultats mitigés de ces initiatives étaient dus au manque d'intérêt de leurs partenaires[174]. Il est vrai que les bénéficiaires étaient souvent méfiants par rapport à l'évaluation, qu'ils voyaient comme une ingérence dans leurs politiques nationales et comme une menace pour leur accès aux futurs fonds d'aide. De plus, ils n'étaient pas formés à l'évaluation des politiques publiques et peu intégrés aux phases de conception et de mise en œuvre des programmes d'aide, ce qui entravait leur capacité à participer aux évaluations par la suite[175]. Les donateurs suédois expliquaient, néanmoins, que si les pays bénéficiaires montraient un intérêt très limité pour les évaluations *ex post*, ils étaient plus réceptifs à une approche pragmatique de l'évaluation, avec des évaluations à mi-parcours, car elles apportaient des résultats qui leur paraissaient plus concrets et opérationnels[176].

De manière générale, nous pouvons conclure que les évaluations participatives étaient considérées, dans les années 1980, comme des bonnes pratiques en théorie, mais peu applicables en réalité car trop coûteuses et trop complexes à mettre en œuvre par rapport à la valeur ajoutée qu'elles apportaient aux donateurs.

---

[172] Entretien avec un évaluateur des Nations unies, 6 octobre 2014.
[173] *Ibid*.
[174] DAC-OECD, *Revue par les pairs : Suède*, 1985.
[175] Ancien agent de coopération français dans les années 1980, entretien mené en mai 2013.
[176] DAC-OECD, 1985, *op. cit.*

- **L'évaluation au service de la rationalisation économique de l'aide**

Plus que par les avancées en matière d'évaluation participative, les décideurs politiques occidentaux étaient très intéressés par les évaluations économiques qui permettaient de mesurer le retour sur investissement des activités d'aide[177]. Nous avons identifié que cette demande provenait du constat généralisé, dans les années 1970, que les capacités limitées « d'absorption » des pays bénéficiaires ne permettaient pas la réussite sur le long terme des projets d'aide. Il devenait donc nécessaire d'effectuer des choix dans la programmation de l'aide au développement, ce qui rejoignait la logique de conditionnalité. Ce fut donc bien une demande politique qui détermina le développement de méthodes d'évaluation économiques et quantitatives plutôt que d'évaluations sociologiques et participatives. La doctrine émergea selon laquelle il ne fallait mettre en œuvre que les projets ayant un effet fort et direct sur la croissance économique des pays du Sud et que ces derniers seraient capables de perpétuer. C'est ainsi que, durant les années 1970, des théories « d'évaluation économique » furent développées par différents organismes d'aide au développement.

Composées principalement d'économistes américains, les équipes de la Banque mondiale développèrent la méthode des « prix de référence » pour mesurer l'efficience économique des projets[178]. Selon cette méthode, les prix observés sur le marché ne traduisaient pas toujours la réalité des échanges économiques. Pour compenser cet écart, il fallait chercher un nouveau système de prix qui traduirait mieux les bénéfices et les coûts réels des projets pour la collectivité. Sur la base de ce nouveau système de prix théorique, il était alors possible de mesurer la rentabilité économique des projets. Il s'agissait donc d'une méthode d'évaluation *ex ante*, mais qui servait de base de référence aux évaluations *ex post* mesurant la valeur ajoutée économique des projets pour les bénéficiaires. Cette méthode d'évaluation économique se diffusa largement dans les agences d'aide anglo-saxonnes et dans les banques de développement.

En France, la suprématie de l'approche économique dans les évaluations était également très visible. Toutefois, les évaluateurs français du ministère de la coopération élaborèrent, dès les années 1960, une méthode d'analyse économique des projets alternative à celle de la Banque mondiale : la « méthode des effets ». Cette équipe d'évaluateurs était principalement composée d'économistes du développement et de polytechniciens, dont le plus reconnu était Marc

[177] Voir bibliographie pour des références précises.
[178] CHERVEL, Marc, « La méthode des effets : présentations récentes, nouvelles citriques », *Tiers-Monde*, tome 22, n°87, 1981, pp. 603-612.

Chervel, spécialiste de la planification qui avait participé durant plusieurs décennies à l'élaboration des principaux guides méthodologiques d'évaluation de l'aide dans les ministères français. La méthode des effets « consiste en une procédure d'analyse et de calcul économique qui vise à mesurer l'intérêt d'un projet du point de vue d'un ensemble national ; dans le ou les calculs, les avantages relatifs au projet sont rapprochés de ses coûts. La méthode des effets relève donc des méthodes dites "coûts-avantages" »[179]. Cette méthode conservait les prix du marché, contrairement à celle de la Banque mondiale, et prenait en compte les bénéfices du projet au-delà du cercle du/des entrepreneur(s) et investisseur(s) direct(s). La méthode des effets calculait les revenus supplémentaires entraînés par le projet, non seulement au niveau « du groupe d'agents » investisseurs et entrepreneurs, mais également au niveau des « agents » de la collectivité nationale dans son ensemble. Cette méthode permit de démontrer que, dans un nombre significatif de cas, les populations les plus aisées bénéficiaient le plus des retombées des projets d'aide car elles étaient plus capables de saisir les opportunités que ces derniers ouvraient, tandis que les populations vulnérables étaient laissées à l'écart.

Bientôt, les projets financés par la France furent, très régulièrement, évalués suivant la méthode des effets. Dans les institutions internationales, cette méthode devint la « méthode française » d'évaluation des projets. Elle se diffusa au Canada, puis dans de nombreux pays d'Europe, d'Asie et d'Amérique Latine et dans les pays bénéficiaires de l'aide française. Le ministère du plan du Sénégal élabora, en 1989, sur la base d'une série d'évaluations de projets agricoles et industriels, un manuel d'évaluation fondé sur la méthode des effets[180]. Cette diffusion de la méthode s'explique, notamment, par le fait que les experts français du ministère de la coopération qui assistèrent le Sénégal et d'autres pays du Sud dans leurs efforts de planification nationale ou sectorielle furent Marc Chervel et son équipe.

L'étude du profil des pays du Sud ayant adopté la méthode des effets, nous a permis d'établir une corrélation entre le succès de cette méthode et le type de discours politiques sur le développement que tenaient les gouvernements. Ils avaient pour point commun de prôner une mise à distance de leur économie par rapport au marché international. Sans chercher une autarcie complète, ils souhaitaient développer des outils pour mettre en œuvre des politiques nationales de développement qui ne seraient pas centrées sur l'aide importée des pays du Nord. Dans un contexte marqué par la libéralisation des échanges, il s'agissait de replacer les politiques de développement dans le contexte national. On voit ici que les préférences des

[179] CHERVEL, Marc, « La méthode des effets trente ans après », *Revue Tiers-Monde,* tome 33, n°132, 1992, pp. 873-891.
[180] CHERVEL, 1992, *art. cit.*

gouvernements en matière d'évaluation reflètent leurs idéologies politiques, ce qui confirme notre thèse selon laquelle l'évaluation est un outil de lecture politique de l'aide au développement.

Les travaux d'AMIRA[181] et de l'INSEE-Coopération travaillèrent également à améliorer la prise en compte des réalités nationales des pays du Sud dans la production d'indicateurs et de cadres d'analyse économique des projets d'aide au développement, en particulier en ce qui concerne le secteur informel. Les travaux d'AMIRA se penchaient sur la manière dont il est possible d'intégrer l'économie informelle dans la mesure du développement des pays du Sud, où elle prend une place très importante. Les travaux d'AMIRA furent innovants et contestataires en montrant que l'orientation des évaluateurs, dans les années 1970, vers le tout quantitatif rendait difficile la compréhension des contextes des pays en développement et des raisons de la réussite ou de l'échec des projets d'aide. De fait, les indicateurs quantitatifs et les données qualitatives se révèlent être davantage complémentaires que substituables[182].

S'il y avait une concurrence méthodologique entre les experts anglo-américains et français en matière d'évaluation, les seconds cherchant à repolitiser, ou à réincarner, l'analyse des projets d'aide au développement, l'objectif restait identique : *ex ante* rationaliser le choix des projets de développement sur des fondements de rendement économique, et *ex post* mesurer les retombées des projets sur la croissance économique des pays du Sud.

Cependant, les membres du groupe de travail sur l'évaluation du CAD-OCDE constatèrent, en 1986[183], que les méthodes d'évaluation économique, aussi bien de la Banque mondiale que française, s'avéraient difficiles à mettre en œuvre dans les pays en développement. Les tutelles gouvernementales des organismes d'aide étaient très demandeuses de ces méthodes, alors que les évaluateurs se heurtaient à d'importantes difficultés méthodologiques : le manque de données économiques nationales, le mauvais suivi des coûts des projets, l'absence d'objectifs d'efficacité fixés en amont des projets, l'incapacité technique à prendre en compte les effets négatifs ou indirects des projets, etc.

[181] Le Groupe AMIRA (1975-1988) est un groupe de recherche informel et pluridisciplinaire pour *l'amélioration des méthodes d'investigation en milieu rural africain*. Y participent des personnes ayant une expérience de terrain et faisant partie de l'INSEE, de l'ex-ORSTOM, de la Coopération technique française, de Bureaux d'études, etc. AMIRA s'oppose à l'utilisation du PIB comme indicateur unique de mesure du développement des pays du Sud et donc de l'efficacité des activités d'aide sur ce développement.
[182] GABAS, Jean-Jacques, « La mesure du développement : comment science et politique se conjuguent », *Revue Tiers Monde*, n°213, janvier 2013, pp. 1-11.
[183] DAC-OCDE, 1986, *op. cit.*

Hormis les articles scientifiques sur la méthode des effets, il existait très peu de travaux académiques sur l'évaluation des projets d'aide au développement en France dans les années 1980, contrairement aux pays anglo-saxons où la littérature sur ce thème proliférait. L'évaluation intéressait peu le milieu de la recherche en sciences sociales en France, bien que le secteur du développement ait été à l'avant-garde en la matière par rapport aux autres politiques publiques. Les méthodes d'évaluation utilisées dans l'aide au développement influencèrent les autres politiques publiques en France, même si cette diffusion fut lente. Ce n'est qu'à la toute fin des années 1980 que la méthode des effets commença à être largement utilisée dans l'analyse économique des projets nationaux[184].

À la fin des années 1980, les évaluateurs des agences d'aide commencèrent à remettre en cause le tropisme économique des évaluations. Les évaluateurs étaient d'autant plus critiques sur ce tropique qu'ils avaient de plus en plus des profils spécialisés en sciences sociales plutôt qu'économiques (excepté à la Banque mondiale) et avaient une compréhension du développement dépassant le facteur de la croissance économique[185].

Les évaluateurs de l'ODA constataient que « les opérateurs des projets ne portent pas assez d'attention à l'étude du contexte et de la faisabilité technique des projets, ainsi qu'au suivi de leur mise en œuvre »[186]. La mauvaise gestion des projets serait la principale cause de leur échec. Basil Cracknell critiquait notamment la conditionnalité des projets de l'ODA, qui donnait trop d'importance aux éléments économiques quantifiables pour sélectionner les projets et ignorait certains aspects fondamentaux, comme leur capacité à s'intégrer dans le contexte social, culturel ou politique des pays du Sud. Selon lui, au lieu de prendre comme indicateur de sélection le taux de retour sur investissement des projets, les décideurs devraient plutôt prendre en compte la capacité des bénéficiaires à gérer et maintenir durablement les projets[187].

Par ailleurs, nous remarquons que les échanges d'expérience sur l'évaluation entre les donateurs étaient très faibles. L'ODA était l'une des seules agences bilatérales à exprimer la volonté de diffuser sa culture et son expérience de l'évaluation au niveau international. Ainsi, elle organisa, en 1986, l'une des premières conférences internationales sur l'évaluation de l'aide au développement. Jusqu'alors, la diffusion d'une culture de l'évaluation de l'aide au développement se réalisait de manière verticale du multilatéral vers le bilatéral. Toutefois, l'influence normative des organismes multilatéraux restait encore faible dans la majorité des

[184] CHERVEL, 1992, *art. cit.*
[185] Voir annexe 10.
[186] DAC-OECD, 1986, *op. cit.*
[187] CRACKNELL, 1984, *art. cit.*

États. L'institutionnalisation de la fonction d'évaluation dans les agences d'aide se faisait donc de manière fragmentée et peu harmonisée.

C. Les prémices de la normalisation et de l'institutionnalisation de l'évaluation

À la fin des années 1970, la Banque mondiale et l'OCDE démarrèrent un travail de normalisation de la fonction de l'évaluation de l'aide au développement. En 1978, le Corps commun d'inspection des Nations unies publia un Glossaire des termes relatifs à l'évaluation. En 1979, le CAD-OCDE conduisit son premier examen comparatif de la qualité de la gestion de l'aide de ses pays membres[188]. Face au constat que la pratique de l'évaluation était peu harmonisée entre les organismes d'aide, le CAD-OCDE créa une plateforme de capitalisation et de normalisation des connaissances et des pratiques d'évaluation. Un groupe *ad hoc* de « correspondants en évaluation » fut créé en 1980, qui avait pour principal objectif de produire des études sur l'efficacité de l'aide à partir des résultats des évaluations réalisées par les donateurs membres. Le groupe devint permanent en 1983 en raison d'une prise de conscience généralisée de l'importance de l'enjeu de la mesure de l'efficacité de l'aide, et donc de l'évaluation, dans une période de crise économique mondiale[189].

- **La faible influence normative des organisations multilatérales**

Au commencement des activités du groupe, à la demande du Secrétariat du CAD, les experts en évaluation réalisaient des sortes de méta-évaluations[190], afin de mesurer l'efficacité macroéconomique de l'aide. Le Secrétariat du CAD souhaitait que ces méta-évaluations servent à renforcer le soutien public en faveur de l'aide au développement. Le groupe d'évaluateurs trouvait cette demande peu réaliste, estimant que « le CAD, en tant qu'entité politique, tendait à avoir des attentes trop ambitieuses sur ce que l'évaluation pouvait apporter. Il était trop optimiste sur le fait que les preuves apportées par les évaluations convaincraient l'opinion publique que l'aide était un bon instrument de développement »[191]. Le groupe d'évaluateurs

[188] DAC-OECD, *Historique du Comité d'aide au développement de l'OCDE,* 2006.
[189] Entretien avec une responsable du département de l'efficacité de l'aide du CAD-OCDE, 2 juillet 2013.
[190] Selon l'OCDE, une méta-évaluation désigne une évaluation conçue comme une synthèse des constatations tirées de plusieurs évaluations. Le terme est également utilisé pour désigner l'évaluation d'une évaluation en vue de juger de sa qualité et/ou d'apprécier la performance des évaluateurs.
[191] DAC-OECD, *Lessons of Experience*, Working group on Aid Evaluation, 1984.

avait conscience que les conclusions des évaluations pouvaient être mitigées, ce qui pourrait potentiellement détériorer l'avis de l'opinion publique sur l'utilité de l'aide. Dans son premier rapport sur l'efficacité de l'aide, en 1982, le groupe souligna, stratégiquement, les limites de son étude. Il reconnut que les évaluations utilisées pour ce rapport ne permettaient pas d'établir de preuves « dures » sur l'efficacité de l'aide. Le groupe utilisa l'opportunité de cet « échec » pour lancer un plaidoyer en faveur d'une redéfinition de son mandat, ce qu'il parvint à obtenir.

En 1984, un rapport du CAD-OCDE stipula que « libéré de la pression de devoir démontrer l'efficacité de l'aide, le groupe peut maintenant se concentrer sur l'apprentissage à partir des leçons de l'expérience passée »[192]. Les évaluateurs avaient donc réussi à modifier la demande politique à leur égard et à gagner une autonomie plus grande dans leurs travaux. Les nouvelles tâches du groupe d'experts étaient de :

- renforcer l'échange et la coopération entre les unités d'évaluation des donateurs ;
- améliorer l'efficacité de l'aide en diffusant des « leçons apprises » sur la conception et la planification des projets ;
- travailler au lancement d'évaluations conjointes entre des membres du CAD-OCDE ;
- promouvoir et soutenir les capacités d'évaluation des pays du Sud.

Ce programme d'activités montre que les membres du groupe d'évaluateurs du CAD-OCDE entendaient avoir un impact normatif fort sur l'évaluation de l'aide. Il convient maintenant de voir si cette ambition normative a pu se concrétiser.

Le groupe d'évaluateurs était constitué des responsables des unités d'évaluation des agences de développement des pays membres du CAD. Dès la première réunion, en 1983, les parties prenantes se sont accordées pour convier le centre de développement de l'OCDE, le PNUD, le FMI, la Banque mondiale et trois banques régionales de développement en tant qu'observateurs. Le groupe a toujours eu à cœur de limiter le nombre de ses membres, tout en trouvant des moyens de collaborer avec d'autres partenaires, afin de préserver sa force, qui tiendrait du fait qu'il soit constitué de pairs. Cette volonté de former un « club » fermé et élitiste est propre aux instances de l'OCDE en général. Des évaluateurs ayant fait partie de ce groupe ont témoigné d'un attachement fort à cette configuration, ce qui posera problème à mesure que le système de l'aide s'élargira à de nouveaux acteurs et se démocratisera[193].

[192] DAC-OECD, 2006, *op. cit.*
[193] Entretien avec un membre du réseau d'évaluateur du CAD-OCDE, 16 novembre 2012 ; entretien avec une responsable du département de l'efficacité de l'aide du CAD-OCDE, 2 juillet 2013 ; entretien avec une analyste des politiques au sein de l'unité d'évaluation du CAD-OCDE, , 2 juillet 2013 ; entretien avec un analyste des politiques au sein du département de l'efficacité de l'aide du CAD-OCDE, 2 juillet 2013.

Les évaluateurs ont également témoigné avoir trouvé en ce réseaux d'expertise un havre de camaraderie salvateur. Selon nous, ce groupe a largement contribué à insuffler un esprit de « communauté » aux évaluateurs du développement. Un évaluateur du groupe expliqua, par exemple, que « travailler avec des collègues qui rencontraient les mêmes problèmes et jouaient des rôles similaires dans leurs pays respectifs a permis de développer une forte camaraderie entre les membres »[194], ou encore, un ancien responsable de l'unité d'évaluation de la DANIDA : « la responsabilité m'ayant été confiée d'établir une unité d'évaluation au sein de l'agence danoise de développement, j'ai appris beaucoup du groupe d'experts. Les membres plus expérimentés partageaient généreusement leurs expériences avec ceux qui étaient nouveaux dans la profession, et nous étions beaucoup. Participer aux réunions était également important pour nous (évaluateurs des agences de développement), car nous étions relativement isolés dans nos organisations. C'était donc un soulagement d'être avec des collègues et de pouvoir partager nos difficultés quotidiennes »[195]. Ce groupe eut une fonction forte de socialisation des évaluateurs qui étaient membres d'une « profession » mal reconnue dans le système de l'aide au développement. À travers ces récits, nous comprenons comment une culture de l'évaluation commença à se diffuser de manière horizontale entre les évaluateurs des agences d'aide, les pays les plus expérimentés en la matière (pays anglo-américains et nordiques) partageant leur savoir-faire avec les pays les moins avancés (France, Belgique, Allemagne, etc.).

En 1986, le groupe d'évaluateurs du CAD-OCDE publia un rapport qui faisait le bilan de l'expérience des donateurs en matière d'évaluation et qui proposait un glossaire, afin d'harmoniser les terminologies utilisées[196]. Ce fut là une étape importante de la normalisation de l'évaluation de l'aide au développement. Néanmoins, les auteurs du glossaire précisèrent en préambule que, même si le document proposait des définitions, la signification des termes dépendait du contexte dans lequel ils étaient utilisés, ainsi que de la culture organisationnelle des agences d'aide. Le groupe soulignait également qu'à mesure que la pratique de l'évaluation se professionnaliserait, des termes nouveaux seraient créés et que les définitions évolueraient.

À cette époque, le CAD-OCDE estimait que la norme de l'évaluation devait se construire, puis se diffuser, naturellement, et non être imposée de manière *top down* d'une organisation multilatérale vers les agences bilatérales. Toute la difficulté pour le CAD était alors de trouver

[194] DAC-OECD, *The DAC Network on Development Evaluation – 30 years of strengthening learning in development*, juin 2013.
[195] *Ibid*.
[196] DAC-OECD, *Méthodes et procédures d'évaluation de l'aide : un recueil des pratiques des donateurs*, 1986.

le juste équilibre entre, d'une part le laissez-faire pour que la norme se diffuse et soit appropriée, mais où chacun risquait d'adopter des conduites divergentes et, d'autre part l'imposition de la norme qui aurait permis qu'elle soit appliquée de manière homogène, le risque étant que son appropriation et sa mise en œuvre ne soient que de façade.

Des initiatives de diffusion de la pratique de l'évaluation et de mutualisation des compétences et des expériences entre évaluateurs se développèrent également en dehors du cadre du CAD-OCDE. L'exemple le plus marquant fut la démarche d'évaluations conjointes[197] de la Commission européenne concernant ses projets de soutien à l'élevage pastoral en Afrique. La méthode des évaluations conjointes fut théorisée pour permettre un apprentissage mutuel et un partage des moyens et des ressources humaines entre les donateurs. De manière générale, la Commission européenne eut une influence forte sur la systématisation de la pratique de l'évaluation des politiques publiques, et plus spécifiquement de l'aide au développement, dans les pays européens, en établissant ses standards propres qui rejoignent ceux du CAD-OCDE[198], en multipliant les évaluation qu'elle menait, et en incitant les donateurs européens à pratiquer l'évaluation. Selon Nicolas Matyjasik, « la Commission européenne a été l'instituteur de l'évaluation en France »[199]. Nous voyons en la mise en œuvre d'évaluations conjointes, un premier signe d'une convergence des approches et des méthodes de l'évaluation entre les pays européens. Toutefois, cette pratique était, et demeure, très rare.

La Banque mondiale eut également un rôle de normalisation de l'évaluation, notamment envers les autres organisations multilatérales. Suivant son modèle, les banques de développement et les agences spécialisées des Nations unies développèrent toutes leurs propres départements d'évaluation durant les années 1980. Les Nations unies et les banques de développement utilisaient des méthodologies et un vocabulaire largement similaires à celui de la Banque mondiale. Cela posa vite problème aux organismes des Nations unies. En effet, le personnel de la Banque mondiale était (et est toujours) très majoritairement issu de formations et de parcours académiques et professionnels économiques et financiers. Leur conception de

[197] Selon le CAD-OCDE, une évaluation conjointe est une évaluation réalisée par plusieurs donateurs qui mutualisent leurs financements et leurs connaissances sur le projet.

[198] Les études sur l'influence de l'union européenne sur l'évaluation des politiques publiques par ses pays membres, notamment en matière d'évaluation étant nombreuses et fouillées, notre thèse revient peu sur cet aspect. Pour davantage d'informations sur ce thème, consulter par exemple : FOUQUET, Annie, MEASSON, Ludovic, *L'évaluation des politiques publiques en Europe : cultures et futurs*, L'Harmattan, 2009, 44p ; JACOB, Steve, *Institutionnaliser l'évaluation des politiques publiques : étude comparée des dispositifs en Belgique, en France, en Suisse et aux Pays-Bas*, PIE-Peter Lang, 2006, 271p ; PERRET, Bernard, 2008, *op. cit.* pp. 110-130 ; BASLE, Maurice, DUPUIS, Jérôme, LE GUYADER, Sylviane, *Évaluation, action publique territoriale et collective*, tome 2, L'Harmattan, 2002, 332p.

[199] MATYJASIK, Nicolas, *L'´évaluation des politiques publiques dans une France décentralisée. Institutions, marché et professionnels*, Thèse pour le Doctorat en Science politique sous la direction de Andy Smith, Université de Bordeaux, 3 décembre 2010, 405p.

l'aide au développement et de son évaluation fut donc influencée par une inclinaison à porter plus d'intérêt aux enjeux économiques du développement qu'à ses enjeux sociaux et politiques ; inclinaison que ne partageaient pas le personnel des Nations unies. En conséquence, en 1976, les Nations unies créèrent la *Joint Inspection Unit* qui eut un rôle important dans le processus de diffusion d'une culture de l'évaluation propre aux agences onusiennes, se distinguant de celle de la Banque mondiale. Son mandat était notamment d'élaborer une terminologie, des standards et une méthodologie de l'évaluation communs aux organisations et agences des Nations unies et de mettre en place des systèmes d'évaluation[200] plus efficaces.

Toutefois, le pouvoir normatif des organismes multilatéraux sur les agences bilatérales restait limité dans les années 1980. En France notamment, cette pression normative fut faiblement ressentie par les évaluateurs. Dans un entretien, un ancien évaluateur du ministère français de la coopération, déclare à ce propos : « *la Banque mondiale était très lointaine de la France et son rayonnement n'était pas celui d'aujourd'hui. En conséquence, la France n'a pas été inspirée par les pratiques de la Banque dans son processus d'institutionnalisation de la fonction d'évaluation des activités d'aide au développement* »[201]. Les évaluations françaises ne se fondaient pas, non plus, sur les normes établies par le CAD-OCDE. Selon un ancien évaluateur du ministère de la coopération, « *il aura fallu attendre la fin des années 1990, soit dix ans après le début de l'effort de normalisation de l'évaluation par le CAD-OCDE, pour que la France adopte les standards de l'organisation* »[202]. En effet, la culture de l'évaluation et de la redevabilité était encore peu présente dans les politiques publiques françaises. Les dirigeants politiques n'incitaient pas les opérateurs de l'aide à intégrer dans leurs stratégies et leurs pratiques les normes internationales en évaluation, et sans ce relai, l'influence normative des organisations multilatérales pouvait difficilement peser.

Les recommandations des examens par les pairs du CAD-OCDE n'avaient que peu d'effets sur les acteurs de l'aide français. Selon un évaluateurs retraité, « *cette réticence était due à la culture administrative française de l'aide au développement, où les acteurs, notamment les dirigeants, étaient peu enclins à l'autocritique* »[203]. Au contraire, lorsque nous interrogeons les évaluateurs britanniques de l'époque, ils témoignent d'une culture forte de la critique et de la prise de recul chez les responsables de l'ODA, qui favorisait la culture de l'évaluation et la

[200] Pour une définition de « système d'évaluation », voir lexique.
[201] Entretien avec un ancien évaluateur du ministère français de la coopération, expert auprès du CAD-OCDE, 21 mars 2013.
[202] Entretien avec un ancien conseiller économique et évaluateur au ministère de la coopération, 6 avril 2013.
[203] *Ibid.*

prise en compte des critiques du CAD-OCDE, lequel était d'ailleurs dominé par la pensée anglo-américaine.

Ces constats nourrissent notre hypothèse selon laquelle la diffusion de « bonnes pratiques » de l'évaluation dépend fortement de la culture politique nationale des pays donateurs et du relai fait par les dirigeants politiques nationaux. Si la culture nationale n'est pas réceptive, l'importation d'une pratique harmonisée de l'évaluation, répondant à des normes qualitatives édictées au niveau international, ne pourra pas prendre. L'introduction de l'évaluation dans les politiques d'aide n'est donc pas simplement un processus de transfert technique, mais demande bien une modification de l'approche politique de l'aide au développement vers plus de transparence, de redevabilité et vers une volonté d'améliorer l'efficacité de l'aide grâce à l'accumulation de nouvelles connaissances, provenant notamment des opérateurs de terrain et des bénéficiaires.

- **Une première phase d'institutionnalisation de l'évaluation dans les agences bilatérales d'aide au développement**

Si l'influence normative des organismes multilatéraux sur les donateurs bilatéraux était variable, les entretiens menés et la littérature étudiée témoignent d'un engouement généralisé pour l'évaluation et un développement de sa pratique dans les années 1980[204]. Malgré la faiblesse de la culture de l'évaluation en France à cette époque, les évaluateurs se remémorent que « *dans les années 1980, on ne parlait que d'évaluation dans le secteur de l'aide au développement* »[205]. Pour beaucoup d'évaluateurs, cette période constitua un « boom » de l'évaluation dans le secteur de l'aide au développement[206].

Dans les années 1980, la plupart des agences d'aide se dotèrent d'une unité spécifiquement dédiée à la fonction d'évaluation, alors qu'auparavant elle était une pratique ponctuellement réalisée par les unités opérationnelles pour apprendre à partir de projets expérimentaux ou pour mesurer le bénéfice de certains projets. L'ODA se dota, dès la fin des années 1970, d'un département de l'évaluation qui rapportait au Sous-Secrétariat chargé de l'aide au développement[207]. Les initiateurs de cette entité, spécifiquement dévolue à

[204] CRACKNELL, 2000, *op. cit,* pp. 44-47; *REBIEN,* 1996, *op. cit.*
[205] Entretien avec un ancien évaluateur au ministère de la coopération français, consultant auprès d'organisations internationales, 21 mars 2013.
[206] CRACKNELL, 2000, *op. cit,* p. 44.
[207] BROWNING, Richard, "Evaluation in the ODA. A view from the inside", *Public Administration and development,* vol. 4, 1984, pp. 133-139.

l'évaluation, étaient motivés par le souci d'engager des personnes qui n'étaient pas directement impliquées dans les projets d'aide et qui pouvaient ainsi avoir un regard critique et indépendant sur les activités évaluées. Cette volonté nouvelle de garantir l'indépendance des évaluations était logique, dans la mesure où l'évaluation avait désormais une fonction d'amélioration de l'efficacité de l'aide et de justification des conditionnalités, questions absente dans les années 1960 et 1970 marquées par une confiance généralisée en les bienfaits de l'aide.

En France, la fonction d'évaluation commença également à s'institutionnaliser dans les organismes de développement : le ministère de la coopération, la direction générale du Trésor et la Caisse française de développement (CFD). D'après un évaluateur retraité du ministère de la coopération, si le ministère ne possédait pas d'unité d'évaluation, il avait tout de même mené une série d'évaluations de 1982 à 1986 qui marquèrent les esprits, en particulier parce qu'elles bousculèrent la vision traditionnelle de l'aide au développement dans les filières agricoles, notamment le coton. L'effet fut d'autant plus important que ce secteur était marqué par une forte idéologie politico-économique, que les évaluations remirent en question. La direction générale du Trésor et la CFD, quant à elles, se dotèrent de structures dédiées à l'évaluation rétrospective des projets au début des années 1980. À la CFD, l'unité d'évaluation était rattachée au Département des politiques et études, ce qui révèle à la fois un souci d'assurer l'indépendance vis-à-vis du personnel opérationnel et une volonté d'établir un lien direct entre les évaluations et les études stratégiques prospectives.

Les deux cas montrent que l'évaluation devint, dans les années 1980, une activité spécifique, demandant un personnel avec des compétences professionnelles adaptées. Toutefois, l'évaluation était encore loin d'être une activité systématique. Les donateurs évaluaient 15 à 20 % de leurs projets, soit 10 à 30 évaluations par an, selon leur volume d'activité. Seule l'USAID évaluait systématiquement ses projets, ce qui équivalait à 250 évaluations par an[208]. Cette spécificité reflète la forte culture de l'évaluation des politiques publiques en Amérique du Nord par rapport aux pays européens. Les évaluateurs européens de l'époque ne considéraient pas nécessaire d'évaluer tous les projets, car cela revenait, selon eux, à réaliser des études superficielles. Au lieu de « saupoudrer », mieux valait se concentrer sur quelques projets novateurs ou présentant des enjeux particulièrement intéressants et mener des analyses plus approfondies[209].

[208] OCDE, 1986, *op. cit.*
[209] CRACKNELL, 2000 *op. cit*, REBIEN, 1996, *op. cit*, ainsi que des témoignages tirés de l'étude de l'OCDE sur les pratiques d'évaluation des agences d'aide en 1986 (voir bibliographie).

Nous notons ici une différence d'approche de l'évaluation entre le niveau politique et le niveau opérationnel. Le niveau politique voyait l'évaluation comme un instrument de mesure des résultats des projets d'aide et poussait à une approche exhaustive de l'évaluation, afin de disposer de l'information la plus large possible sur les performances de l'aide, alors que les opérateurs et la plupart des évaluateurs restaient dans une logique d'apprentissage et dans une approche sélective de l'évaluation, se concentrant sur les activités pouvant apporter une connaissance nouvelle à l'organisation.

Nous allons désormais analyser les effets de cette divergence de perception sur la pratique de l'évaluation.

Les dirigeants des agences d'aide essayaient de mêler une approche « sommative » des évaluations, pour satisfaire la demande politique de leurs tutelles, et une approche « formative », pour répondre à la demande des opérationnels. Ils mirent ainsi en place des équipes mixtes d'évaluation, c'est-à-dire qui incluaient des évaluateurs externes et internes à l'organisation. Les évaluateurs externes permettaient une prise de recul et une indépendance, nécessaires à une approche « sommative » de l'évaluation. Les évaluateurs internes, quant à eux, dans une logique formative, garantissaient la pertinence des évaluations pour l'apprentissage interne et aidaient à l'acceptation et à la diffusion des recommandations des évaluations au sein de l'organisation.

Concernant l'ODA, le personnel de l'agence estimait qu'il était très important de continuer à intégrer les évaluateurs de l'organisme à la mise en œuvre des évaluations[210]. Cela permettait un meilleur apprentissage interne des leçons tirées et une meilleure ouverture du personnel aux nouvelles idées avancées et aux éventuelles critiques prononcées. Certains allaient même jusqu'à demander un retour à une mise en œuvre complètement internalisée des évaluations, comme c'était le cas durant les décennies précédentes. Le recours à des évaluateurs externes était considéré comme une obligation des tutelles financières et politiques de l'ODA et du CAD-OCDE, qui cherchaient à garantir l'indépendance des évaluations, les utilisant dans un objectif sommatif. Le personnel de l'ODA était très méfiant à l'égard des évaluateurs externes[211]. Les évaluateurs internes et les opérationnels de l'agence critiquaient, par exemple, la longueur des rapports d'évaluation fournis, alors qu'ils aimeraient des rapports concis et faciles à lire, exposant clairement les enjeux et les leçons des projets évalués, au lieu de « les noyer dans une masse de considérations méthodologiques »[212]. Ces critiques des évaluateurs

[210] CRACKNELL, Basil, 1986, *art. cit,* BROWNING, 1984, *art. cit.*
[211] BROWNING, 1984, *art. cit.*
[212] BROWNING, 1984, *art. cit.*

internes témoignent aussi d'un sentiment de concurrence avec les évaluateurs externes qui menaçaient leur fonction.

Une autre critique courante à l'encontre des évaluateurs externes consistait à dire qu'ils avaient une compréhension approximative de la culture organisationnelle de l'ODA. En conséquence, les recommandations émises par les évaluateurs externes auraient souvent été inappropriées. Les agents de l'ODA soulignaient aussi que les rapports ne traitaient pas suffisamment des enjeux qui les intéressaient et des leçons qui seraient utiles à l'organisation pour le futur. Un membre de l'ODA a des propos illustratifs à ce sujet : « *les évaluateurs passent souvent à côté de l'objectif essentiel qui est de répondre à la question suivante : comment pouvons-nous, en tant qu'organisme d'aide, mieux faire les choses ? Et non pas comment les choses pourraient être mieux faites en général ?* »[213].

Ces témoignages confirment que les recommandations d'une évaluation seront d'autant plus acceptées et utilisées si les parties prenantes concernées sont associées à l'exercice. Les rapports humains et sociaux ont une grande importance dans la réussite d'un processus évaluatif. Le personnel de l'ODA réclamait le droit d'effectuer un cadrage et d'avoir un contrôle sur les travaux menés par des évaluateurs externes. Cette demande et les critiques précédentes pouvaient être le reflet, soit d'une réelle mauvaise qualité des rapports d'évaluation, soit de la difficulté du personnel à se voir évalué par des intervenants extérieurs, ou encore d'une tentative d'ingérence dans le travail des évaluateurs pour obtenir des rapports favorables. Nous estimons que ces trois aspects étaient mêlés dans le comportement des agents de l'ODA vis-à-vis des travaux des évaluateurs externes.

En France, la situation était plus diversifiée. Les évaluations n'étaient pas systématiquement faites par des équipes mixtes ou externalisées. Selon des évaluateurs de la CFD, les évaluations étaient réalisées en interne avec l'appui ponctuel de consultants extérieurs spécialisés sur des enjeux de développement spécifiques[214]. À la direction générale du Trésor, les équipes d'évaluation étaient mixtes et leurs travaux s'orientaient vers la mesure de la rentabilité économique des projets pour la France. Certains évaluateurs d'autres organismes de développement critiquaient les évaluations que menait le Trésor et les qualifiaient de « travaux d'évaluation d'une grande hypocrisie »[215], dans la mesure où ils auraient été biaisés pour mettre

[213] Entretien avec un évaluateur interne de l'ODA, 14 janvier 2013.
[214] Élément recueilli durant une observation participante à l'AFD en 2012 et 2013 qui a permis un dialogue continu avec les évaluateurs et agents de l'AFD sur leur expérience de l'évaluation.
[215] Entretien avec un ancien évaluateur au ministère de la coopération français, consultant auprès d'organisations internationales, 4 avril 2013.

en valeur les résultats positifs de la France dans les pays du Sud. Cette critique laissait percevoir une tension entre les différents organismes d'aide sur la détermination des enjeux à traiter dans les évaluations et les méthodes à adopter. Une sorte de compétition qualitative s'installa entre les unités d'évaluation des organismes d'aide français. Le ministère de la coopération, quant à lui, externalisa assez rapidement l'intégralité des évaluations. Nous expliquons cette décision par le fait que le ministère était déstabilisé et menacé de disparition et qu'il avait besoin de légitimation externe. Dans cette organisation « en crise », l'évaluation avait donc un rôle de légitimation plus marquée et nécessitait davantage le recours à des évaluateurs externes. On peut en déduire que plus la fonction de légitimation est mise en avant dans les objectifs des évaluations, plus les organismes font appel à une expertise externe.

Lorsque nous étudions les documents stratégiques des organisations de développement de l'époque et que nous interrogeons leur personnel[216], nous remarquons qu'être évaluateur n'était pas vraiment un métier, mais plutôt une fonction parmi d'autres au sein du métier d'expert thématique sur le développement, de chercheur ou de chargé de projets. Les évaluateurs n'étaient donc pas des acteurs en tant que tels dans le système de l'aide au développement, leur statut n'étant pas reconnu. Ainsi, il était commun de voir des bureaux d'études externes mener à la fois les études de faisabilité, de planification et d'évaluation des projets d'aide.

Toutefois, malgré cette assimilation des métiers, nous avons noté qu'il y avait une tendance généralisée des agences d'aide à instituer l'évaluation comme une fonction à part entière avec la création d'unités d'évaluation, même si ce changement n'était pas encore assimilé par le personnel, ni même par la plupart des évaluateurs eux-mêmes. La fonction d'évaluation n'étant pas encore bien déterminée, les unités d'évaluation se « *baladaient de départements en départements au sein des agences* »[217], passant du département stratégique à celui du contrôle interne, ou encore de l'audit. L'évaluation cherchait donc sa place dans les organismes d'aide, où l'environnement n'était généralement pas favorable à son essor.

En France, la perception de la fonction d'évaluation dans les organismes d'aide au développement était assez mauvaise. Les évaluateurs et les agents de l'époque tiennent des propos révélateurs à ce sujet : « *l'unité d'évaluation était souvent un placard sympathique où l'on mettait des diplomates ou des agents un peu marginaux* »[218]. Dans les années 1980,

[216] Archives internes de l'AFD et entretiens avec des agents de l'AFD et de l'ODA.
[217] Entretien avec un ancien évaluateur au ministère de la coopération et expert en évaluation auprès de l'OCDE, 21 mars 2013.
[218] Entretien avec un évaluateur de l'AFD, 14 juin 2013.

l'évaluation était vue comme un métier de généraliste qui ne nécessitait pas de compétences techniques spécifiques. Les évaluateurs souffraient donc d'un manque de légitimité et d'autorité pour faire accepter leur expertise. D'ailleurs, certaines personnes interrogées vont jusqu'à sous-entendre que si les décideurs choisissaient de nommer des « agents marginaux » à cette fonction, c'était pour qu'ils n'aient pas assez de poids pour imposer des changements importants sur la politique et la gouvernance des organisations[219]. En conséquence, peu d'agents étaient volontaires pour être placés dans les unités d'évaluation, car cela n'était pas bénéfique pour leur carrière.

Tout comme au Royaume-Uni, les unités opérationnelles étaient méfiantes vis-à-vis de l'expertise des évaluateurs, d'autant plus s'ils étaient externes aux organisations. Un ancien agent du ministère de la coopération nous a raconté qu' « *au départ, il a fallu forcer la main aux services opérationnels pour qu'ils acceptent de se faire évaluer et de prendre en compte les recommandations des évaluations. Il a fallu beaucoup de temps pour que les services acceptent le regard extérieur de l'évaluateur. Si les évaluations pouvaient permettre un apprentissage en interne, le discours qui concernait la redevabilité relevait de la logorrhée* »[220].

Cette étude de l'arrivée de la fonction d'évaluation dans les agences d'aide, qui demanda au personnel d'avoir un regard critique sur leurs activités, montre que l'introduction d'une nouvelle fonction dans une structure organisationnelle demande un temps relativement long d'acceptation. Nous voyons également que la mise en place et l'utilisation d'évaluations à des fins de redevabilité eut du mal à s'inscrire dans la logique des opérateurs de l'aide et des évaluateurs, peu habitués à rendre comptes des résultats de leurs activités.

Les évaluations étaient peu utilisées par les décideurs politiques[221]. Les évaluateurs témoignent que la reconnaissance de l'importance de l'évaluation ne correspondit pas à une augmentation de leur utilisation, ni même de leur mise en œuvre[222].

En France, selon un ancien évaluateur du ministère de la coopération, qui se remémore son expérience professionnelle des années 1980, « *la prise en compte des recommandations par les décideurs était faible. Les évaluations avaient fait apparaître beaucoup de dysfonctionnements de gestion dans les organismes d'aide au développement : manque de coordination entre les services en interne et entre les différents organismes chargés de la*

[219] Entretien avec un évaluateur retraité de l'AFD, 14 septembre 2013.
[220] Entretien avec un ancien évaluateur au ministère de la coopération français, consultant auprès d'organisations internationales, 4 avril 2013.
[221] Entretien avec un évaluateur retraité de l'AFD, 14 septembre 2013.
[222] RIDDEL, Roger. C, *Foreign Aid Reconsidered,* The John Hopkins University Press, 1987, 320p.

coopération, manque d'organisation, manque de cohérence, etc. Ces constats et les recommandations qui en découlaient n'étaient pas pris en compte par les décideurs »[223]. En effet, dans les années 1980, les évaluateurs n'avaient pas assez d'autorité et de légitimité pour pouvoir faire valoir leurs recommandations auprès des dirigeants.

Au-delà de la question du statut des évaluateurs, la faible utilisation des évaluations relevait aussi de leur nature et de leur qualité. Les évaluations se focalisaient sur l'échelle des projets et se contentaient souvent de mesurer leurs réalisations dans une logique d'analyse coûts-bénéfices. Plus rarement, elles s'intéressaient aux processus de mise en œuvre des projets dans une logique d'apprentissage. La nature des évaluations n'était donc pas adaptée aux intérêts des décideurs politiques. Pour intéresser ces derniers, il aurait fallu qu'elles prennent en compte les résultats plus globaux de l'aide, à l'échelle des secteurs d'activité ou des pays. Les évaluateurs étaient conscients de cette demande des décideurs, mais la complexité méthodologique de l'évaluation globale de l'aide a empêché, pendant un temps, le développement de telles évaluations.

En 1986, Robert Cassen, économiste à l'université d'Oxford, dirigea une étude retentissante dans le milieu de l'évaluation, qui démontra que les rapports d'évaluation ne constituaient pas des preuves satisfaisantes de l'efficacité de l'aide, car ils se concentraient davantage sur la bonne allocation des fonds et des moyens que sur les résultats finaux des projets[224]. Pour que l'évaluation soit utile, il fallait qu'elle porte sur les résultats des programmes d'aide et juge de leur efficacité. C'est de preuves d'efficacité qu'avaient besoin des dirigeants. Néanmoins, la prise de conscience que l'efficacité de l'aide devait être évaluée ne fut pas accompagnée d'efforts importants pour développer la théorie et la méthodologie de l'évaluation. Les experts en évaluation de l'époque regrettaient que peu de budget soit consacré à la recherche méthodologique en évaluation[225]. C'était surtout vrai dans les pays donateurs non anglophones, où la littérature grise et la recherche sur l'évaluation, et plus spécifiquement sur l'évaluation de l'aide au développement, étaient très maigres.

Pour contrer le défaut d'utilisation des évaluations, l'ODA se dota, en 1987, d'un *Projects and Evaluation Committe* composé de représentants de la direction de l'organisation. Il avait pour tâche de contrôler la qualité des évaluations et d'assurer que les recommandations émises étaient bien prises en compte par le ministère chargé du développement. Ce comité constituait

[223] Entretien avec un ancien évaluateur au ministère de la coopération français, consultant auprès d'organisations internationales, 4 avril 2013.
[224] CASSEN, Robert, *Does Aid Work? Report to an international task force,* Clarendon press, 1986, 381p.
[225] CRACKNELL, 1988, *art. cit.*

un mécanisme d'incitation auprès du personnel pour qu'il utilise les évaluations. L'ODA décida également de porter davantage d'attention à la bonne rédaction des termes de référence des évaluations et de produire un manuel de « bonnes pratiques » d'évaluation à destination des évaluateurs externes[226] pour assurer la qualité de leurs rapports. Ainsi, la fonction d'évaluation se formalisa au sein de l'agence. Dans les années 1980, la culture de l'évaluation était déjà forte à l'ODA. Cet ancrage de l'évaluation s'explique par l'importance donnée à l'évaluation dans les politiques publiques au Royaume-Uni et par l'influence des pratiques américaines et de la Banque mondiale sur l'agence. L'ODA était donc sous une double incitation, aux niveaux national et international, pour systématiser sa pratique de l'évaluation.

Au contraire, les organismes français étaient encore loin d'avoir comme priorité l'harmonisation des procédures et de la qualité des évaluations, qui étaient produites par trois entités séparées, échangeant peu sur leurs méthodes et leurs productions respectives, et entretenant même une certaine compétition. Concernant les mécanismes de suivi de la prise en compte des recommandations des évaluations, il n'en existait pas dans les organismes français de développement. L'utilisation des évaluations dépendait donc de l'intérêt que portaient les acteurs concernés à ce type d'analyses.

- **Conclusion de chapitre**

Ce premier chapitre nous a permis d'observer le « déterminisme politique » qui marque les évolutions de la théorie et de la pratique de l'évaluation. La manière dont les évaluations sont menées reflète des conceptions particulières de l'aide au développement.

Dans les années 1960 et 1970, les donateurs adoptaient une approche tiers-mondiste du développement dans le Sud et étaient sûrs de l'efficacité de leurs pratiques d'aide. Ils avaient alors peu recours à l'évaluation et l'utilisaient uniquement à des fins d'apprentissage à propos de projets expérimentaux ou à des fins de mesure du ratio « coûts-bénéfices » des projets d'aide.

En 1980, l'aide au développement prit une tournure néolibérale et devint conditionnelle. L'efficacité de l'aide commençait à être remise en cause par le milieu de la recherche. Les donateurs voulurent alors évaluer les résultats de l'aide et utiliser l'évaluation pour rationaliser leurs allocations vers les pays les plus stables et sûrs. Nous voyons ainsi que l'évaluation était

[226] CRACKNELL, 1988, *art. cit.*

politisée et s'adaptait aux besoins des acteurs de l'aide en y conformant ses objets d'étude et sa méthodologie.

De plus, ce chapitre a permis de montrer que la diffusion de standards d'évaluation du niveau multilatéral au niveau bilatéral dépend fortement de l'existence d'un « relai » des dirigeants politiques et d'une culture préexistence de l'évaluation des politiques publiques au niveau national. Ainsi, alors qu'au Royaume-Uni, l'influence normative des organismes multilatéraux se faisait ressentir, dans un contexte où ces derniers étaient dominés par la pensée anglo-américaine du développement et que l'évaluation des politiques publiques gagnait en importance dans les administrations britanniques, cette influence ne se fit pas ressentir en France, où la culture de l'évaluation des politiques publiques était moindre et où les organismes d'aide se sentaient éloignés des organismes multilatéraux. Ainsi, les pratiques d'évaluation étaient disparates entre les agences bilatérales et multilatérales d'aide au développement. Chaque agence d'aide évaluait selon sa culture, ses besoins et ses contraintes organisationnelles.

Nous allons désormais voir comment cette diversité des approches a évolué avec l'intensification de l'effort de normalisation de la fonction d'évaluation des organismes multilatéraux et la diffusion internationale des approches anglo-américaines des politiques publiques dans les années 1990.

Chapitre 2 : 1991-2000, l'évaluation, un outil de légitimation de l'aide au développement

La période que nous allons étudier s'étend de 1991 à 2000. Nous avons choisi ces dates car elles marquent des étapes importantes dans le processus de diffusion d'une culture de l'évaluation dans le secteur de l'aide au développement. 1991 est la date où le CAD-OCDE a adopté un document normatif, Les critères de qualité pour l'évaluation de l'aide au développement, qui va devenir le référentiel international de « bonnes pratiques » en matière d'évaluation. Nous allons montrer que les années 1990 furent la période où l'évaluation s'est normée et institutionnalisée dans les organismes internationaux. Cette norme de l'évaluation a été construite par les représentants des unités d'évaluation des donateurs membres du CAD-OCDE et sous l'influence de l'expertise de la Banque mondiale. Les plus actifs dans ce processus étaient les représentants des pays anglo-saxons pour lesquels la gestion des politiques publiques était marquée depuis les années 1980 par un phénomène de « bureaucratisation néolibérale »[227] qui favorisa l'essor de la demande d'évaluation par les parlements et les gouvernements.

La décennie 1990 fut marquée par plusieurs tendances qui modifièrent le rapport des acteurs de l'aide à l'évaluation. La méthode gestionnaire du New Public Management se diffusa dans les pays de l'OCDE et transforma les pratiques de gestion et d'évaluation de l'aide. Elles se focalisèrent alors sur des enjeux de transparence et de redevabilité, enjeux qui intéressaient peu les acteurs de l'aide auparavant. Cette demande de transparence fut exacerbée par un sentiment généralisé de « crise de l'aide » suite à l'échec de la politique du Consensus de Washington. La tendance amorcée dans les années 1980 s'affirma donc. Pour répondre à cette demande, les organismes d'aide se mirent à la recherche d'instruments leur permettant de légitimer leurs actions auprès de leurs tutelles financières et de l'opinion publique.

À la fin de la décennie, des experts du développement, notamment au PNUD, inspirés par les travaux d'Amartya Sen[228], cherchèrent à sortir d'une vision purement économiste des processus de développement pour donner davantage d'importance aux aspects sociaux, environnementaux et institutionnels. Cela changea la temporalité de l'aide au développement

[227] HIBOU, Béatrice, 2012, *op. cit*. Le terme de bureaucratisation néolibérale renvoie notamment, selon Béatrice Hibou, à un phénomène de remplacement de la hiérarchie dans les organismes publics par des procédures de contrôle indirectes, telles que l'évaluation, l'audit, etc.
[228] SEN, Amartya, *On ethics and Economics*, Basil Blackwell, 1987, 192p.

qui fut considérée comme un accompagnement de long terme des États et non comme un soutien temporaire. Au sein des organismes d'aide, l'analyse des anthropologues fut davantage considérée, comme en témoigne Michael Cernea, premier « sociologue-anthropologue » à avoir rejoint la Banque mondiale en 19874[229]. Si les effets de l'aide sur l'économie pouvaient révéler leur efficacité sur le court terme (bien qu'ils puissent être réversibles), les effets sur l'environnement, les structures sociales ou institutionnelles ne pouvaient se concrétiser que sur le long terme, demandant des transformations plus profondes de la culture et des modes de vie des sociétés. Ce changement entraîna une complexification de la fonction d'évaluation, qui portait désormais sur des enjeux nouveaux et plus qualitatifs, où la mesure économétrique ne suffisait plus aux évaluateurs pour juger de l'efficacité de l'aide. Les évaluateurs durent donc développer de nouvelles méthodes d'analyse face à une demande qui évoluait rapidement.

Une tension apparut alors entre une gestion de l'aide qui s'inspirait du New Public Management (Cf. infra) avec, d'une part des exigences de résultats à court terme et d'évaluation constante de la performance et, d'autre part une politique de l'aide qui s'orientait vers une approche institutionnelle et sociale du développement demandant des actions de long terme dont les résultats « immatériels » étaient difficilement quantifiables.

Dans les années 1990, la pratique de l'évaluation augmenta, se complexifia et se normalisa. Pour être une institution d'aide au développement légitime, il fallait prouver son efficacité et donc évaluer ses performances et les rendre publics. Une logique de comparaison internationale des performances des organismes d'aide se mit alors en place, qui nécessitait que les évaluations soient harmonisées (pour être comparables). Ce phénomène s'accentua dans les années 2000[230].

Ainsi, les organismes internationaux d'aide au développement – Banque mondiale, Nations unies, CAD-OCDE - entreprirent un effort de normalisation de la fonction d'évaluation. Cette dernière devint une norme qu'il s'agissait de respecter pour apparaître comme un bon gestionnaire aux yeux de la « communauté internationale du développement ». Ces organismes internationaux firent peser une pression normative sur les agences d'aide bilatérales pour qu'elles adoptent des « bonnes pratiques » en matière d'évaluation. Nous verrons alors comment les agences d'aide ont réformé leurs structures internes pour se conformer à ces « bonnes pratiques » et comme elles se les sont appropriées.

[229] CERNEA, Michael. M, *la Dimension humaine dans les projets de développement*, Karthala, 1999, 586p, p. 44.
[230] *Cf.* Chapitre 3 et 9.

A. L'évaluation comme élément de réponse à la perte de légitimité de l'aide

Dès le début des années 1990, les donateurs prirent conscience que la stratégie du Consensus de Washington des années 1980 avait été un échec. En conséquence, la légitimité de l'aide, jusqu'alors forte, fut détériorée. L'activité d'évaluation pris alors de l'ampleur dans les organismes d'aide, comme un instrument permettant de prouver l'efficacité de l'aide au développement.

- **L'échec du Consensus de Washington et ses conséquences sur la fonction d'évaluation**

Une recherche bibliographique, nous a permis de décompter plus de 4 000 livres ou articles scientifiques consacrés au sujet de l'efficacité de l'aide au développement depuis les années 1960 avec un essor de 1985 à 1990 et après 1995[231]. Nous avons parcouru un certain nombre de ces publications, ce qui nous a amené à remettre en cause leur valeur ajoutée pour la recherche sur l'aide au développement. En effet, plus qu'un débat sur les fondements politiques et économiques sous-jacents à la réussite des projets d'aide au développement, ils constituaient une « guerre des écoles » entre auteurs, qui confrontaient leurs méthodes de mesure de l'efficacité de l'aide (micro-économique, macro-économique, qualitative, quantitative, etc.), entraînant une perte de « sens » de ces travaux. Il s'agissait plus d'un débat méthodologique sur la manière la plus adéquate de mesurer l'efficacité de l'aide que d'une réflexion pour améliorer les pratiques de l'aide. Certains de ces travaux, que nous allons présenter ci-dessous, ont toutefois marqué l'histoire de l'évaluation de l'aide au développement.

Au fil des ans, un nombre de plus en plus important de ces publications remettaient en cause l'efficacité de l'aide[232]. À ces travaux répondaient des études qui réhabilitaient la capacité de l'aide à entraîner des processus vertueux de développement dans les pays bénéficiaires. Parmi les chercheurs dont les travaux sont les plus cités, nous retrouvons Paul Mosley, chercheur à l'University of Cambridge et l'University of Sheffield au Royaume-Uni et économiste au DFID et à la Banque mondiale[233], ainsi que Robert Cassen, chercheur à la LSE

[231] Recherche à partir de Ngram Viewer : https://books.google.com/ngrams/graph?content=aid+effectiveness&year_start=1960&year_end=2008&corpus=15&smoothing=3&share=&direct_url=t1%3B%2Caid%20effectiveness%3B%2Cc0, consulté le 10 novembre 2014.
[232] Pour une liste de ces travaux, se référer à l'annexe 3 qui présente l'état de l'art de la littérature sur l'efficacité de l'aide au développement.
[233] MOSLEY, 1980, *op. cit.*

et à l'Oxford University et expert sur le développement humain au DFID et à la Banque mondiale[234], ou encore Craig Burnside, chercheur à la Quinn University et à l'University of Pittsburg aux États-Unis et expert économiste à la Banque mondiale, et enfin, David Dollar[235], chercheur à l'UCLA aux États-Unis et économiste à la Banque mondiale. Ces chercheurs regroupent deux caractéristiques communes : ils sont rattachés à des laboratoires de recherche anglo-américains tout en ayant des activités d'expertise au sein des organisations de développement les plus influentes. Nous voyons ici un moteur essentiel de l'influence de leurs travaux sur la pensée du développement des années 1990. Ce double statut leur permit d'accéder aux données des organismes de développement, d'observer comment ils fonctionnaient de l'intérieur, tout en conservant une approche analytique et critique. Leurs travaux s'apparentaient à des « méta évaluations », c'est-à-dire des études qui fondent leurs conclusions sur l'agrégation et l'analyse de plusieurs évaluations menées par des organismes d'aide au développement. .

Le retentissement des travaux remettant en cause l'efficacité de l'aide dans l'opinion publique, *via* les ONG et la presse, entraîna des controverses sur l'utilité de l'aide au développement. Ces travaux confortaient le sentiment de plus en plus généralisé que l'aide ne parvenait pas à lutter contre la pauvreté. Les organismes d'aide se virent alors contraints de prouver l'utilité de leurs activités sur la base de preuves rigoureuses, sous peine de perdre leur légitimité d'action dans les pays du Sud. Ce devoir de rendre des comptes fut renforcé par une baisse massive des budgets d'aide au développement dans les années 1990[236] qui obligea les organismes d'aide à rationaliser davantage l'allocation de leurs fonds. Une analyse à partir des statistiques du CAD-OCDE montre que le volume d'aide au développement passa de 0,32 % du PIB des pays de l'OCDE en 1992 à 0,21 % en 1997[237]. Cette baisse des fonds et l'affaiblissement de la légitimité de l'aide ont entraîné une demande accrue d'évaluation de la performance des projets et programmes d'aide au développement. Nous observons ici une corrélation entre l'accentuation de la crise de l'aide et la hausse de la demande en évaluation à des fins de redevabilité, au détriment des objectifs d'apprentissage.

Ce devoir de redevabilité fut un changement important pour les acteurs de l'aide. Dans les années 1960 et 1970, l'aide se justifiait d'elle-même par des principes moraux. Si l'efficacité de l'aide posait déjà question dans les années 1980, c'était avec une ampleur bien moindre que

[234] CASSEN, 1986, *op. cit.*
[235] BURNSIDE, Craig, DOLLAR, David "Aid spurs growth – in a sound policy environment", *Finance and Development*, vol. 34, n°4, Washington, 1997, pp. 4-7.
[236] Baisse généralisée par rapport aux années 1980 où elle ne fut que ponctuelle et toucha un nombre plus restreint d'États.
[237] NAUDET, Jean-David, 2005, *op. cit.*

dans la décennie 1990, la confiance en le Consensus de Washington étant encore forte. C'est ainsi que John Lewis, président du CAD, déclarait dans un discours de 1980 qu'« il n'y a pas plus lieu de demander une preuve de l'efficacité de l'aide que de demander une preuve de l'efficacité de l'impôt »[238]. Nous n'excluons pas que l'efficacité ait été un enjeu de l'aide au développement avant les années 1990, mais cet enjeu était borné au niveau de la bonne affectation des fonds et des réalisations de l'aide, sans que soit vraiment abordée la question de l'amélioration générale du niveau de développement des pays bénéficiaires. De plus, cette évaluation globale de l'efficacité de l'aide au niveau des États ou sur l'ensemble des pays en développement paraissait impossible à réaliser par les évaluateurs et ne faisait pas l'objet de demandes spécifiques de la part des commanditaires des évaluations (décideurs politiques et dirigeants des agences d'aide) qui auraient poussé les évaluateurs à tenter de développer des méthodes adaptées à ce type d'analyses. Par ailleurs, le contexte de Guerre froide apportait une légitimité stratégique aux activités d'aide au développement, justifiées par la volonté de maintenir les pays du Sud dans le Bloc de l'Ouest.

Au Royaume-Uni, des documents politiques[239] illustrent cette évolution des fondements de l'aide en parlant, par exemple, *d'enlightened legislation* ou de *long-term self interest*[240] pour justifier l'aide au développement. Selon cette approche, l'aide au développement n'était pas que de la charité, elle répondait également aux intérêts de la nation. C'est dans les années 1990 qu'une approche « conséquentialiste »[241] des activités d'aide s'institua, selon laquelle les bons résultats de l'aide conditionnaient sa légitimité. Le principe d'efficacité devint alors un fondement de l'aide au développement et l'évaluation un outil central permettant de mesurer les résultats de l'aide et donc d'en prouver l'efficacité et la légitimité. Ainsi, l'évaluation fut considérée comme une fonction plus stratégique au sein du système de l'aide au développement, que les acteurs voulurent maîtriser pour en tirer des bénéfices en termes de légitimité.

Devant la mise en cause de l'efficacité des anciennes pratiques d'aide au développement, les bailleurs abandonnèrent l'idéologie du Consensus de Washington fondée sur le principe de conditionnalité. Les experts de la Banque mondiale proposèrent, comme alternative, ce qu'ils

---

[238] *Ibid.*

[239] Secretary of State for International Development, *White Paper on International Development: Reducing poverty a challenge for the 21st century*, novembre 1997.

[240] Ces termes rappellent le concept d' « enlightened self interest » issu de la philosophie et qui exprime l'idée qu'une entité qui agit pour améliorer la condition d'une ou de plusieurs entités tierces va ultimement servir son propre intérêt. Il se rapproche du concept de « main invisible » développé en économie et permet de conjuguer une approche éthique et altruiste de l'aide à une approche rationnelle de type gagnant-gagnant. L'aide au développement apparaît donc légitime tant du point de vue des bénéficiaires que des pays donateurs et de leurs contribuables.

[241] Terme repris à Jean-David Naudet dans ses travaux sur l'évolution des paradigmes de l'aide au développement cités en bibliographie.

présentaient comme une méthode d'allocation « optimale » de l'aide faisant la synthèse entre des indicateurs de sélection centrés sur « le besoin », c'est-à-dire le niveau de pauvreté des États, et des indicateurs de bonne gouvernance[242], prenant en compte la qualité des institutions[243].

Selon certains experts de l'aide au développement, s'inscrivant souvent dans des courants anthropologiques ou sociologiques[244], cette approche centrée sur les notions d'efficacité, de résultat, de redevabilité et de sélectivité fut porteuse de conséquences négatives sur les fondements de l'aide au développement. Le changement affiché vers une aide plus juste est aujourd'hui souvent considéré comme étant une illusion. C'est l'une des thèses développées par Jean-David Naudet, ancien directeur du département de l'évaluation à l'AFD, « le retentissement de la problématique d'efficacité de l'aide n'a pas été sans provoquer des dérives téléologiques. La confusion entre le juste et le performant est manifeste, notamment dans la doctrine de la sélectivité. En effet, si l'on quitte la conditionnalité au sens du Consensus de Washington, c'est pour en retrouver une nouvelle, rebaptisée « sélectivité », qui se fonde sur des principes de bonne gouvernance et une perspective de réduction d'une pauvreté importante en amont des projets »[245]. Si la sélectivité se voulait plus juste, car fondée sur un principe méritocratique – les pays ayant les meilleures perspectives de performance recevaient le plus d'aide –, elle favorisait en fait les pays souffrant du moins de handicaps, notamment en termes de bonne gouvernance[246]. Les populations des pays les « moins bien dirigés » se retrouvaient donc défavorisées par rapport à celles de pays « bons élèves », dans la même logique que le Consensus de Washington.

Nous pensons que si la sélectivité apparut comme un renouveau et séduisit tellement les bailleurs, c'est parce qu'elle donnait une justification rationnelle à leurs stratégies d'allocation discriminatoire de l'aide. La sélectivité se basait sur des indicateurs quantitatifs qui donnaient une impression de neutralité. Elle apparaissait fondée, non pas sur la condition réformes imposées par les donateurs, mais sur des facteurs purement internes aux pays bénéficiaires de l'aide : leur niveau de pauvreté, la qualité de leur gouvernance, leur capacité à mener des

[242] Ces indicateurs qui servent à évaluer la bonne gouvernance des pays du Sud et leurs effets seront étudiés en détail dans le chapitre 8 de la thèse.
[243] COLLIER, Paul, DOLLAR, David, *Aid Allocation and Poverty Reduction*, World Bank, Development Research Group, avril 1999.
[244] Voir par exemple LECOMTE, Bernard, NAUDET, Jean-David, « Survivre grâce à l'aide, réussir malgré l'aide », *Cahier des Sciences humaines*, n°13, Éditions de l'Aube, 2000, 211p ; RIST, Gilbert, *Le développement, Histoire d'une croyance occidentale*, Presses de Sciences-po, 2013, 511p.
[245] NAUDET, Jean-David, 2005, *op. cit.*
[246] Pour une définition du terme « bonne gouvernance » voir lexique.

réformes. La sélectivité devint ainsi la nouvelle bonne pratique promue par la Banque mondiale afin de permettre une allocation rationnelle et optimale de l'aide.

Cette bonne pratique favorisa la hausse de la demande en évaluation. En effet, les planificateurs des activités d'aide ont besoin d'évaluations *ex ante* de l'état de développement des pays pour fixer des objectifs atteignables, et d'évaluations *ex post* systématiques pour mesurer les résultats des projets d'aide par rapport aux objectifs fixés.

- **La réduction de la pauvreté et la bonne gouvernance, nouveaux paradigmes centraux de l'aide**

La crise de l'aide ressentie au début des années 1990 ne relevait pas seulement d'une perte de légitimité pouvant être résolue par un effort en termes de transparence et de redevabilité pour rassurer les tutelles financières, les cercles académiques et l'opinion publique. Cette crise était aussi interne aux organismes d'aide au développement. Face au creusement des inégalités dans le monde, que l'aide au développement ne parvenait pas à résorber, les acteurs du développement souffraient d'un sentiment de désillusion. Dans les sommets internationaux, la possibilité de parvenir à un état de développement égalitaire entre le Nord et le Sud n'était plus évoquée[247]. Les objectifs de l'aide au développement étaient revus à la baisse. Ils se concentraient sur la réduction de la pauvreté la plus extrême, objectif qui apparaissait plus atteignable que l'égalisation des conditions entre le Nord et le Sud. L'OCDE, la Banque mondiale et les Nations unies firent alors de la « lutte contre la pauvreté » le nouvel objectif universel de l'aide auquel adhérèrent tous les donateurs. Antoine Pouillieute, ancien directeur général de l'AFD, a témoigné de ce mouvement global : « l'émergence du thème de la lutte contre la pauvreté dans la communauté internationale n'est pas le fait de l'un ou l'autre des acteurs de la coopération, mais le fait d'une réflexion collective qui s'est structurée par étapes. Il y a eu, tout d'abord, les critiques faites au fameux Consensus de Washington et à ses politiques libérales, avec, notamment, le rapport intitulé *l'Ajustement à visage humain de l'Unicef,* en 1987, puis la montée en puissance de la théorie du développement humain durable. Les pays de l'OCDE, emmenés par les pays du Nord, réfléchissent concrètement à des directives pour réduire la pauvreté. Enfin, la crise asiatique de 1997-1998 a définitivement fait vaciller le Consensus de Washington »[248].

[247] CRACKNELL, 2000, *op. cit.* p. 48.
[248] POUILLIEUTE, Antoine, « la lutte contre la pauvreté, nouveau paradigme de l'Agence française de développement », *Politique africaine,* n°81, 2001, pp. 171-175.

En 1996, l'OCDE publia un rapport affirmant que « l'aide au développement exprime le sentiment de compassion que suscitent l'extrême pauvreté et les souffrances humaines qui affligent encore aujourd'hui le cinquième de la population mondiale »[249]. Les organisations multilatérales incitèrent les acteurs de l'aide à centrer tous leurs efforts sur la réduction de la pauvreté. Pour les guider dans cette voie, elles élaborèrent des documents normatifs exposant les bonnes pratiques à adopter pour lutter efficacement contre la pauvreté[250]. Le recentrage des politiques d'aide autour de cet objectif plus modeste, et surtout à forte valeur morale, visait à persuader les donateurs bilatéraux d'augmenter leurs volumes d'APD et favorisa un regain de légitimité de l'aide. Par ailleurs, la lutte contre la pauvreté apparaissait comme un moyen de sécuriser l'espace mondial. Le lien entre pauvreté et insécurité, qui montera en puissance dans les années 2000, préfigurait déjà clairement dans le discours de l'OCDE : « la pauvreté et la misère qui existent dans le monde menacent la sécurité de tout un chacun »[251]. Dans cette affirmation, on retrouve implicitement l'idée selon laquelle la pauvreté serait un enjeu global contre lequel le Nord comme le Sud devraient s'engager. La sécurité alimentaire devint alors un « bien public mondial »[252].

La mise sur agenda de la réduction de la pauvreté fut donc un instrument de (re)légitimation de l'aide au développement. Si la baisse de la pauvreté avait toujours été une finalité de l'aide, nous percevons comme une nouveauté la place centrale qu'elle prit alors par rapport aux autres thématiques. En témoigne le nombre de publications des organismes d'aide qui avaient pour thème principal la réduction de la pauvreté. La Banque mondiale, par exemple, prit pour thématique la pauvreté pour son *Rapport annuel sur le développement dans le monde* en 1990. De même pour le PNUD, qui consacra, la même année, *son Rapport sur le développement humain à la pauvreté*, ou encore pour le Royaume-Uni qui en fit le pilier de sa nouvelle politique d'aide présentée dans un livre blanc de 1997 intitulé *Réduire la pauvreté dans le monde : le défi du 21ème siècle*[253]. Cette place donnée à la pauvreté était en rupture avec les politiques néolibérales d'ajustement structurel du Consensus de Washington. Toutefois,

[249] CAD-OCDE, 1996, *op. cit.*
[250] DAC-OCDE, *Shaping the 21st Century: The Contribution of Development Cooperation*, mai 1996, 24p; World Bank, « Poverty » in *World Development Report 1990*, 1990.
[251] CAD-OCDE, 1996, *op. cit.*
[252] Selon une définition du PNUD un bien public mondial est un bien ou service dont les caractéristiques inhérentes en termes de non-exclusion et de non-rivalité (voir encadré) rendent difficile l'établissement de droits de propriété ou d'usage, excluant ainsi la possibilité de rentabiliser sur un marché les efforts de production ou de conservation. Quand ils sont produits gratuitement par la nature, ils sont d'ailleurs souvent menacés par la surexploitation : même si l'intérêt collectif est de les préserver, la meilleure stratégie individuelle est de se servir avant que le voisin ne le fasse aussi. Il est donc nécessaire pour produire et protéger ces biens de se doter de règles et de moyens collectifs, coordonnés par une communauté locale ou par la puissance publique nationale.
[253] Secretariat for International Development, 1997, *op. cit.*

nous considérons que ce discours sur la pauvreté cachait une perpétuation des politiques des décennies précédentes, les instruments de l'aide restant les mêmes sous d'autres dénominations.

À mesure que les organisations internationales concentrèrent leurs réflexions sur l'analyse des causes de la pauvreté et des moyens d'en sortir, la notion de pauvreté se complexifia. Dans les années 1990, la Banque mondiale réaffirma une conception économique de la mesure de la pauvreté, selon laquelle cette dernière proviendrait d'un phénomène d'exclusion du marché du travail international. En conséquence, les programmes d'aide devaient mener des actions pour y réintégrer les pauvres. Cette exclusion du marché aurait souvent été due à une mauvaise gouvernance de l'État qui aurait ruiné la confiance des investisseurs internationaux et augmenté la fragilité de l'appareil économique national face aux chocs externes. Les projets sur la santé, l'éducation ou la bonne gouvernance, développés par la Banque mondiale dans le cadre de la lutte contre la pauvreté, étaient considérés comme des investissements dans le capital humain que constituaient les pauvres.

Nous identifions l'« individualisation » de la vision du développement et de la pauvreté comme un des changements majeurs du régime de l'aide dans les années 1990 ; changement notamment porté par la Banque mondiale. L'aide de la Banque se mit à viser les individus plutôt que les États. La cause principale donnée à la pauvreté était la mauvaise gouvernance et les populations pauvres étaient davantage présentées comme des victimes que comme des acteurs de leur pauvreté. Il s'agissait de contourner l'État et de réduire la place du secteur public. L'aide au développement devait donc permettre à ces individus de recouvrer leur potentiel d'enrichissement, d'augmenter leur niveau de vie et par conséquent le niveau de développement de leur État. Nous estimons que l'emphase mise sur les individus, « les pauvres », sous une allure empathique, manifestait leur volonté de responsabiliser « les pauvres » face à leur condition, perpétuant ainsi la logique néolibérale de la décennie précédente. Paradoxalement, si les populations pauvres n'étaient pas « la cause » de leur état de pauvreté, elles étaient censées avoir la capacité de dépasser cet état en adoptant les comportements recommandés par les donateurs.

De leur côté, à partir des travaux d'Amartya Sen[254], les Nations unies, et plus spécifiquement le PNUD, développèrent une approche normative de la pauvreté, définie comme un manque d'opportunités et de liberté. Dans son *Rapport sur le développement humain* de 1997, le PNUD décrivait la pauvreté comme « la négation des capacités humaines

[254] SEN, Amartya, *Development as freedom*, Oxford University Press, 1999, 384p.

fondamentales – analphabétisme, malnutrition, longévité réduite, mauvaise santé maternelle, etc. - qui engendre l'impossibilité pour la personne de réaliser ses potentialités ou de saisir les opportunités dont elle dispose »[255]. La notion de pauvreté dépasserait donc un manque de revenus et intègrerait la qualité de vie générale des individus et leur potentiel d'épanouissement dans un environnement donné.

Ainsi, les actions de lutte contre la pauvreté prirent l'aspect de programmes multisectoriels de plus grande ampleur qui intégraient des dimensions éducative, de santé, de renforcement des capacités institutionnelles, etc. Cette approche globale de la lutte contre la pauvreté suivait la théorie du *big push* développée dans les années 1940 par Paul Rosenstein-Rodan, selon qui la lutte contre la pauvreté passerait par des investissements massifs et simultanés dans tous les secteurs des politiques publiques[256]. Cette approche « globale » de la pauvreté est donc opposée à la stratégie « d'individualisation » de la Banque mondiale.

En ligne avec son approche néolibérale du développement, la Banque mondiale reprit le concept de gouvernance qui émergea dans les années 1970 et 1980 pour qualifier la régulation du système mondialisé. Elle lui donna une nature normative en le transformant en « bonne gouvernance » incarnée par la doctrine de retrait de l'État et d'émancipation économique des acteurs privés. Cette bonne gouvernance fut prônée dans les États du Sud et devint une condition de l'allocation des fonds d'aide. Toutefois, dès la fin des années des années 1990 avec la crise financière asiatique les premiers doutes s'emparent des acteurs de l'aide sur l'efficacité de la recette néolibérale de développement. Des économistes, tels que Jospeh Stiglitz, remirent en cause cette formule[257] et la Banque mondiale produit un rapport transformant son approche du développement intitulé *The State in a Changing World*[258] qui prônait un retour encadré de l'État dans la régulation économique.

[255] UNDP, *Rapport mondial sur le développement humain : Le développement humain au service de l'éradication de la pauvreté*, New 1997.
[256] ROSENSTEIN-RODAN, Paul, "The International Development of Economically Backward Areas", *International Affairs*, vol. 20, n°2, avril 1944, pp. 157-165.
[257] Voir par exemple : STIGLITZ, Joseph, *State versus Market: Have Asian currency crises affected the reform debate*, Published for TheBangladesh Economic Association 50p.
[258] World Bank, *World Development Report 1997 : The State in a Changing World*, 1997.

B. Élargissement et complexification de la fonction d'évaluation

Nous avons émis l'hypothèse que ces nouvelles approches du développement développées par les organismes internationaux eurent un impact fort sur le travail des évaluateurs. Nous allons désormais vérifier cette hypothèse en analysant les influences à partir desquelles les évaluateurs ont fait évoluer leurs pratiques.

- **Le défi de l'évaluation des facteurs sociaux et institutionnels de la pauvreté**

La Banque mondiale avait adopté de longue date un indicateur normatif de mesure de la pauvreté : le « seuil de pauvreté absolue »[259]. Cet indicateur économique de pauvreté servait communément de référentiel aux évaluateurs pour juger de la performance des projets d'aide au développement en matière de lutte contre la pauvreté. En conséquence de son approche « micro » de la pauvreté, la Banque mondiale incita les évaluateurs des organismes d'aide à évaluer les résultats de l'aide au niveau des individus et lier ces résultats à l'observation d'un développement plus global. Dans les séminaires internationaux sur l'évaluation de l'aide au développement, l'adoption d'une approche micro de l'évaluation était encouragée : « la clé de l'approche « quantum » de l'évaluation est que chaque petit changement, même au niveau de l'individu ou du foyer, importe pour le système entier au niveau sectoriel, du pays ou international. Nous devons réaliser les évaluations de manière systémique »[260]. Ainsi, cette approche entraîna la systématisation des évaluations au niveau des projets et préfigura l'essor des évaluations expérimentales quelques années après.

Dans une étude réalisée avec Jean-Jacques Gabas, nous observons que cette mesure économique de la pauvreté ne correspondait plus, dans les années 1990, à l'approche de la pauvreté développée par les Nations unies, qui mettait en avant des critères d'éducation, de santé et d'égalité en plus de celui de revenu monétaire[261]. Les Nations unies développèrent alors des indicateurs plus qualitatifs qui prenaient notamment en compte les écarts de pauvreté internes aux pays et l'intensité de la pauvreté. Ainsi, au début des années 1990, les indicateurs de pauvreté se multiplièrent : le *headcount ratio* désignait la part de la population qui vivait

[259] Seuil de 1 $ par jour en parité des pouvoirs d'achat, qui fut ensuite augmenté à 2$ par jour pour mieux correspondre au contexte des pays à revenus intermédiaires.
[260] Alison, Evans, Séminaire de la Banque mondiale sur l'évaluation des programmes de lutte contre la pauvreté, archives de la Banque mondiale, 2000.
[261] GABAS Jean-Jacques, LAPORTE, Camille, 2011, *art. cit.*

sous le seuil de pauvreté, le *poverty gap* représentait l'écart de niveau de vie de la population pauvre par rapport au seuil de pauvreté et le *squared poverty gap* mesurait l'inégalité des revenus dans la population pauvre. Nous relativisons l'évolution qu'apportèrent ces indicateurs sur l'évaluation de l'efficacité des programmes d'aide en termes de lutte contre la pauvreté. Bien qu'ils constituent une avancée en représentant les inégalités au sein des sociétés, ces indicateurs restaient centrés sur une mesure de la pauvreté en termes de revenus.

C'est la création de l'indice de développement humain (IDH) par le PNUD, en 1990, qui renouvela la mesure de la pauvreté, et en conséquence les critères d'évaluation de l'efficacité de l'aide. Cet indice mesure le développement des États en fonction de trois aspects : l'espérance de vie à la naissance, le niveau d'instruction et le niveau de vie, exprimé par le PIB par habitant. En 1997, le PNUD élabora un indicateur complémentaire de mesure de la pauvreté spécifiquement, *l'indicateur de pauvreté humaine* (l'IPH), qui analysait la pauvreté en termes de manque sans intégrer le PIB. La construction de cet indice par le PNUD montrait la volonté de l'organisation de compléter, voire de dépasser, les standards de la Banque mondiale.

À mesure que la notion de pauvreté s'enrichissait avec des éléments qualitatifs, l'évaluation des programmes de lutte contre la pauvreté se complexifia. Un évaluateur britannique témoigne : « *évaluer la contribution de programmes à la réduction de la pauvreté est un exercice intrinsèquement difficile, bien plus difficile que les évaluations classiques de projets. En effet, la réduction de la pauvreté est complexe et implique beaucoup d'acteurs et de processus sociaux et économiques qui ne sont pas faciles à contrôler, où même à influencer. Une autre difficulté est que les équipes d'évaluation n'incluent généralement pas de représentants des populations pauvres* »[262]. Il s'agissait d'agréger des indicateurs de performance à la fois économiques (PIB, seuil de pauvreté), sociaux (égalités hommes-femmes, éducation, santé) et politiques (bonne gouvernance) et de les pondérer entre eux afin d'évaluer les progrès de la lutte contre la pauvreté. Encore fallait-il disposer de bases de données statistiques complètes sur plusieurs années, ce qui était rare dans les pays en développement où l'appareil statistique était peu développé.

Les évaluateurs durent donc mobiliser des compétences et des savoirs multiples sur le développement, qui leur étaient souvent étrangers. En conséquence, ils s'associaient régulièrement avec des chercheurs spécialisés sur le genre, la santé, le développement durable, etc. En effet, dans les années 1990 au Royaume-Uni, la demande politique d'évaluations

[262] Entretien avec un ancien évaluateur interne de l'ODA, 14 janvier 2013.

participatives était très faible. On demandait aux évaluateurs d'examiner les effets des projets de développement sur la réduction de la pauvreté, sans leur donner le mandat et les moyens d'interroger les populations pauvres sur les changements que les projets avaient apportés sur leurs conditions de vie.

Comme nous l'avons vu, plus du niveau de pauvreté, la bonne gouvernance des États était l'un des facteurs principaux guidant la sélectivité de l'aide et était vue comme une condition de son efficacité. Il était donc primordial pour les évaluateurs de développer des méthodes pour juger des apports des programmes d'aide sur cet aspect. En conséquence, une méthodologie de l'évaluation du renforcement des capacités institutionnelles fut développée par les évaluateurs, notamment de la Banque mondiale, à la fin des années 1990. Ils produisirent des guides établissant des séries d'indicateurs pour évaluer la gouvernance des États. Basil Cracknell, ancien évaluateur à la Banque mondiale, se souvient qu'« à la Banque mondiale, toutes les évaluations devaient obligatoirement intégrer la durabilité et le développement institutionnel, qui étaient considérés comme des critères majeurs de succès ou d'échec des projets »[263].

Un autre aspect du développement que les évaluateurs devaient désormais prendre en compte était l'impact environnemental. Dans les décennies précédentes, l'enjeu environnemental était secondaire. Il était courant que des projets d'exploitation forestière ou de grandes infrastructures endommagent gravement des forêts ou des cours d'eau, bouleversant les modes de vie des populations environnantes. Cette situation entraîna des critiques virulentes de la part des ONG internationales de protection de l'environnement qui poussèrent les organismes d'aide à développer des bonnes pratiques en la matière. Nos entretiens auprès d'organismes de développement[264] nous ont montré que la crainte des critiques de la société civile internationale est toujours un moteur puissant poussant les bailleurs de fonds à renforcer leurs mesures de suivi-évaluation de la protection de l'environnement[265]. Nous pouvons par exemple citer le cas des projets de grands barrages, pour lesquels les évaluations *ex ante*, à mi-parcours et *ex post* sur les effets environnementaux furent, dès lors, exhaustives. Toutefois, en dehors des projets où l'enjeu de la durabilité était très prégnant, comme les grands barrages, l'évaluation des effets environnementaux n'était pas systématique dans les années 1990[266].

[263] CRACKNELL, 2000, *op. cit,* p. 249.
[264] Pour des raisons de confidentialité, nous ne pouvons pas donner de plus amples détails sur le contenu des études que nous avons pu mener durant cette observation participante.
[265] Voir liste des entretiens menés au Laos, annexe 8.
[266] *Ibid*.

Concernant le genre, des études menées par les organismes d'aide au développement démontrèrent que les femmes avaient tendance à être plus gravement touchées que les hommes par la pauvreté, en raison notamment de l'inégalité du partage des ressources (de revenus, d'éducation, d'accès à la terre) en leur défaveur[267]. Les femmes furent donc spécifiquement ciblées par un ensemble de programmes de lutte contre la pauvreté. Ces études montrèrent également que les femmes jouaient un rôle fondamental dans la réduction de la pauvreté, dans la mesure où elles avaient une activité économique informelle importante. En conséquence, les évaluateurs firent face à une nouvelle demande de mesure des effets des projets sur les femmes. Cela constitua pour eux un défi important cat ils avaient souvent peu de connaissances sur la thématique du genre et éprouvaient des difficultés à évaluer le phénomène informel de la contribution des femmes à l'économie locale, régionale ou nationale. Selon Basil Cracknell, « si le nombre d'évaluations prenant en compte le genre augmente, elles restent d'un niveau exploratoire et la marge de manœuvre pour leur amélioration est grande. Plusieurs facteurs entravent une évaluation en profondeur des projets en termes de genre : le manque d'expertise au Nord comme au Sud, le manque d'intérêt des acteurs de l'aide et des évaluateurs pour cette thématique, le manque de statistiques agrégées par sexe, l'absence d'objectifs en matière de genre déterminés en amont des projets, etc. En conséquence, le traitement du genre dans les évaluations reste superficiel : par exemple, les évaluateurs rendent compte du nombre de femmes ayant participé à des formations plutôt que de l'amélioration de l'accès à l'emploi pour les femmes grâce à ces formations »[268].

Face à la faiblesse de l'expertise sur le genre et à la difficulté des agences d'aide à intégrer la notion de genre dans leurs projets, l'OCDE, les Nations unies et la Banque mondiale engagèrent un travail important de sensibilisation sur ce sujet et montrèrent en exemple leur engagement pour la cause de l'égalité homme-femme[269]. Notre recherche a montré une influence forte des organismes multilatéraux sur la prise en compte du genre dans les programmes et les évaluations des organismes bilatéraux de développement. L'OCDE condamna, par exemple, le retard de la France sur cette question dans une revue par les pairs[270]. En conséquence, selon des membres de l'organisme que nous avons interrogés, la CFD recruta Geneviève Verdhelan-Cayre, spécialiste des problématiques de l'égalité homme-femme, pour

[267] UNDP, 1997, *op. cit.*
[268] CRACNELL, 2000, *op. cit.* pp. 251-254.
[269] En 1995, James Wolfensohn, alors président de la Banque mondiale choisit de donner le premier discours de son mandat à la quatrième conférence mondiale sur les femmes à Pékin, où il proclama l'objectif de l'éducation primaire universelle des garçons et des filles d'ici à 2010. En 1999, le CAD-OCDE adopte des « lignes directrices pour l'égalité homme-femme et le renforcement du pouvoir des femmes dans le cadre de la coopération pour le développement ».
[270] DAC-OECD, *Development Cooperation Review: France. L'élimination de la pauvreté dans le monde : faire de la mondialisation un succès pour les pauvres,* 2000.

qu'elle devienne « responsable des questions de genre ». Ce poste, situé dans le département de la stratégie, donna lieu à plusieurs chantiers : des expériences-pilotes sur l'intégration du genre dans différents projets, des formations, ou encore la création d'un groupe de travail « genre » à la CFD[271]. Toutefois, la prise en compte du genre dans les évaluations n'était pas un chantier de réflexion. Par exemple, lorsque le projet d'assistance technique à la Fédération des paysans du Fouta Djallon en Guinée, qui faisait du genre un enjeu central, fut évalué, cette problématique ne fut pas prise en compte. Le rapport de 79 pages ne mentionne pas une seule fois les mots « genre » ou « femme »[272]. Autrement dit, la réflexion qui avait été effectuée en amont sur le genre ne réussit pas à atteindre la phase finale de l'évaluation. Les évaluateurs témoignaient ne pas avoir les connaissances et les instruments nécessaires pour évaluer la prise en compte du genre et les résultats des projets sur le genre. Selon Geneviève Verdhelan-Caye, la faiblesse de la prise en compte du dans les évaluations de projets était liée au fait que « l'instruction des projets était dominée par une logique économique et financière qui laisse peu de place à l'analyse socio-économique et donc à l'analyse de genre. La difficulté se situe également dans le suivi des impacts et des bénéficiaires »[273]. Furent aussi pointées la « culture institutionnelle » de la CFD et celle des partenaires et des pays d'intervention, qui auraient fait barrage à une bonne prise en compte du genre.

Nous percevons, à travers cet exemple, comment la culture organisationnelle des agences d'aide peut entraver la diffusion de bonnes pratiques internationales malgré une pression forte des donateurs. Pour que cette diffusion soit effective, il faut que ces bonnes pratiques soient intériorisées, ou appropriées, par les décisionnaires de l'action publique au niveau national. En témoigne le regain d'intérêt de l'AFD pour le genre et l'environnement dans les années 2010, à la suite d'une demande directe de son ministre de tutelle, Pascal Canfin, qui s'affaiblit lorsque ce dernier quitta son poste.

Au Royaume-Uni, le gouvernement mit en place une large réforme de sa politique d'APD pour réorienter ses objectifs vers l'élimination de l'extrême pauvreté. Dans deux livres blancs datant de 1997 et 2000, le gouvernement affirma son engagement à travailler pour la lutte contre la pauvreté dans les pays en développement[274]. Ces textes encadraient un ensemble de réformes

[271] CORTINOVIS, Mathilde, *Engagements et résistances au genre dans le développement : le cas de l'Agence française de Développement, 1997-2013*, sous la direction de la Professeure Delphine Dulong, Master Coopération internationale, Action humanitaire et Politiques de Développement, Université Paris 1, décembre 2013, 78p.
[272]AFD, *Bilan de l'assistance technique à la Fédération des paysans du Fouta Djallon (FPFD) en Guinée : 15 ans d'accompagnement*, Division Évaluation et capitalisation, Edition Ex Post, n°32, août 2010.
[273]VERDELHAN-CAYRE, Geneviève, *L'AFD et la prise en compte du genre*, note opérationnelle, AFD, janvier 2001.
[274]Secretary of State for International Development, 1997, *op. cit.*; Secretary of State for International Development, *White Paper on International Development: Eliminating World Poverty: Making Globalization Works for the Poor*, 2000.

mises en place parallèlement à l'examen par les pairs du CAD-OCDE en 1998[275] sur l'APD du pays. Ils faisaient explicitement référence aux normes que le CAD-OCDE avait fixées dans son rapport *Shaping the 21st Century*, ce qui tend à montrer le poids normatif fort qu'avait alors acquis le Comité. Le gouvernement mit l'accent sur l'importance de fixer et d'atteindre des cibles mesurables de réduction de la pauvreté, dont la principale était de la diminuer de moitié d'ici à 2015, objectif établi par les Nations unies au niveau international. L'objectif de la réduction de la pauvreté et les indicateurs liés en termes d'égalité homme-femme, d'environnement, de bonne gouvernance, etc. fixés par le gouvernement, servaient de cadre aux évaluateurs pour intégrer les thèmes du genre, de l'environnement et de la bonne gouvernance dans leurs travaux, appui qui a manqué aux évaluateurs français. En effet, si la lutte contre la pauvreté devint également le nouveau paradigme central de l'aide au développement française à cette époque, il manquait les indicateurs et les cibles pour en mesurer les progrès[276].

Selon nous, l'inclusion, plus ou moins effective, des problématiques de la durabilité, du genre et de la bonne gouvernance dans les pratiques d'évaluation des agences d'aide bilatérales se fit grâce à la pression des organisations multilatérales et de la société. Elle fut facilitée dans les pays où la culture des politiques publiques était plus sensibilisée à ces problématiques. Dans les années 1990, les évaluateurs les intégraient de manière expérimentale et tâtonnante dans leurs travaux. Comme nous le verrons dans le chapitre suivant, une sorte d'émulation se créa entre les équipes d'évaluation des différentes agences d'aide dans les années 2000 pour développer la meilleure expertise possible sur ces thèmes. Cette émulation fit de ces enjeux des thématiques centrales de l'évaluation de l'aide, symboles de pratiques exemplaires et sophistiquées. De plus, l'élargissement des thématiques de l'évaluation conduisit les agences d'aide à recruter des évaluateurs internes et des consultants au profils de plus en plus diversifiés et davantage spécialisés dans les sciences sociales et environnementales qu'en économie[277].

- **Combiner expertise scientifique et instrument participatif**

Face aux défis méthodologiques évoqués ci-dessus, les évaluateurs firent le constat de deux manques principaux dans les évaluations : elles ne faisaient pas assez participer les divers acteurs des projets et programmes d'aide et elles avaient une vision à court terme et étroite des effets de l'aide[278]. Les évaluateurs se trouvaient confrontés à une demande paradoxale des

[275] DAC-OECD, *Development Cooperation Review: United Kingdom*, 1998.
[276] POUILLIEUTE, Antoine, 2001, *art. cit.*
[277] Voir annexe 10.
[278] REBIEN, 1996, *op. cit.* pp. 81-90.

donateurs exigeant, d'une part que les évaluations traitent d'enjeux de plus en plus larges, touchant à des thématiques pour lesquelles les bases de données chiffrées étaient rares dans les pays en développement et, d'autre part que les évaluations utilisent des méthodes économétriques jugées plus fiables que les méthodes qualitatives et permettant d'analyser l'efficacité de l'aide au niveau micro sur les individus.

Face au scepticisme des pays donateurs quant aux bénéfices des approches qualitatives de l'évaluation, les évaluateurs développèrent des argumentaires pour les réhabiliter. Par exemple, lors d'une conférence à la Banque mondiale en 1999, plusieurs évaluateurs concentrèrent leurs interventions sur la promotion de ce type d'évaluations : « la nature complexe de la plupart des objectifs du développement requiert des méthodologies calibrées, appropriées et praticables sur le terrain. Étant donnée la complexité de la tâche des évaluateurs et l'ampleur des contraintes opérationnelles dans les pays en développement, une approche moins ambitieuse de l'évaluation peut parfois mener à de meilleurs résultats (…). La qualité de l'analyse est plus importante que la quantité. Une approche modeste de la méthodologie de l'évaluation, qui laisse plus de place au qualitatif qu'au quantitatif, nous apparaît également la plus réaliste »[279]. Nous notons ici une remise en cause de la tendance qui était en cours de systématisation de l'évaluation à tous les projets d'aide.

La nécessité de développer des méthodologies qualitatives et participatives de l'évaluation, déjà présente dans les discours des évaluateurs dans les années 1980, fut largement relayée auprès des décideurs nationaux et internationaux de l'aide au développement. Les évaluateurs défendaient l'idée que de ces méthodologies contribueraient à relégitimer l'aide au développement en marquant une rupture avec les pratiques *top down* du Consensus de Washington. Ils argumentaient qu'elles permettraient de conduire à une autonomisation des populations pauvres et à une démocratisation de la gouvernance, redirigeant les bénéfices des changements sociaux vers les plus pauvres[280], ce qui était en ligne avec l'accent mis par les décideurs des organismes d'aide sur l'individualisation du développement, la bonne gouvernance et la réduction de la pauvreté.

Cette vision de l'évaluation participative relevait de l'utopie. Pour que les évaluations aient pu avoir un tel pouvoir émancipateur sur les populations les plus pauvres, il aurait fallu qu'elles aient lieu dans des contextes politiques favorables, où les gouvernements mettent en

[279] World Bank, *Evaluation and Poverty Reduction,* proceedings of a World Bank Conference, OED, Washington DC, 1999.
[280] ROBB, Caroline, *Can the Poor Influence Policy? Participatory Poverty Assessment in Developing World,* World Bank, 1999.

place des réformes pour améliorer le caractère participatif de la gouvernance et où les programmes d'aide sont planifiés et mis en œuvre de manière participative. De plus, il aurait fallu que les évaluations soient systématiquement participatives et utilisées par les décideurs politiques du Sud. Or, les évaluations des programmes d'aide avaient peu d'écho dans les pays bénéficiaires[281].

Les principaux bénéfices de l'utilisation de méthodes qualitatives et participatives dans les évaluations se limitaient à une meilleure compréhension par les donateurs du contexte social des projets et des attentes et de l'opinion des bénéficiaires. Malgré le plaidoyer des évaluateurs, les responsables favorisaient toujours les méthodes quantitatives, permettant une mesure chiffrée de la performance. À ce sujet, Elizabeth Withmore, chercheuse et consultante auprès d'agences de développement, déclare : « le problème est que nous opérons dans un système qui tend à demander des résultats mesurables sur le court terme. Même lorsque les administrateurs des agences d'aide sont favorables à des approches participatives, ils doivent mobiliser des efforts de persuasion importants pour convaincre les dirigeants d'en autoriser la mise en œuvre »[282]. Selon nous, cette exigence de mesure des résultats par les dirigeants était exacerbée par le contexte des années 1990, marqué par une volonté de relégitimer l'aide au développement. Pour la classe politique, de bons chiffres étaient de meilleurs outils de communication sur l'efficacité de l'aide que des analyses qualitatives difficilement synthétisables.

Un rapport de force apparût alors entre les opérateurs de l'aide au développement et leurs tutelles administratives et financières sur le format et le contenu que devaient adopter les évaluations. Les détenteurs du pouvoir financier et décisionnaire étant les tutelles, l'issue de ce rapport de force pencha en faveur du maintien d'une approche majoritairement quantitative des évaluations.

Ainsi, malgré une prolifération des manuels et guides de bonnes pratiques sur l'évaluation participative dans les agences d'aide[283], leur application restait inexorablement marginale. En 1997, le Conseil scientifique de l'évaluation en France relevait que « l'évaluation de l'aide au développement demeure fortement axée sur le projet ou le programme, donne un

[281] Pour un développement plus approfondie sur le renforcement d'une culture de l'évaluation dans les pays du Sud, voir chapitre 5.
[282] WHITMORE, Elizabeth, "Understanding and Practicing Participatory Evaluation", *New Directions for Evaluation*, n° 80, 1998.
[283] Par exemple : UNDP, *Who are Question-makers ? A Participatory Evaluation Handbook*, Office of Evaluation and Strategic Planning, 1997; World Bank, *World Bank Participation Sourcebook*, Operations Evaluation Department, 1996; Institute for Development, Studies, *Participatory Monitoring and Evaluation: An Introductory Pack*, University of Sussex, 1997.

rôle extrêmement faible aux autorités locales, s'organise autour de missions de terrain brèves et assurées par des consultants extérieurs à qui on ne donne ni le temps ni les moyens d'enquêter sur le terrain »[284].

En outre, la volonté des évaluateurs de mobiliser des outils relevant des méthodes de recherche utilisées par les anthropologues pour rendre leurs évaluations participatives entraîna une réaction négative de ces derniers. Les anthropologues se méfiaient des évaluateurs qui empiétaient sur leurs terrains d'expertise. Cette concurrence se fit notamment ressentir par les critiques que portaient les anthropologues sur le caractère participatif des études menées par les évaluateurs. En France par exemple, certains anthropologues qualifiaient les évaluations participatives « d'exercices imposés par les bailleurs, aboutissant à des diagnostics peu significatifs et à des actions standardisées »[285]. Ils soulignaient également le risque que le caractère participatif des évaluations serve *a posteriori* de caution à des projets qui se déroulaient en réalité de manière *top down*, c'est-à-dire où les donateurs imposaient leurs conditions et leurs méthodes aux bénéficiaires. Parmi les biais récurrents des méthodes participatives, ils identifiaient « le manque de liberté dans la prise de parole dû à la présence de notables (durant les consultations avec les populations bénéficiaires) et à la croyance des évaluateurs qu'un grand nombre de participants garantit, en soi, le caractère participatif d'une évaluation »[286]. Le statut des évaluateurs étant peu reconnu dans le système de l'aide, ils durent faire de gros efforts de plaidoyer pour faire valoir leur expertise par rapport à celle des anthropologues qui les dénigraient[287]. Nous voyons ici comment l'arrivée d'une nouvelle « fonction » dans un secteur d'activité perturbe les acteurs déjà « installés » et crée des rapports conflictuels le temps que chacun se (re)positionne.

Un autre défi posé aux évaluateurs par les dirigeants des agences d'aide et les responsables politiques fut d'inscrire les évaluations dans une perspective de long terme. En effet, les résultats des projets en matière de bonne gouvernance, de protection de l'environnement et de genre ne pouvaient s'apprécier que dans le temps. En outre, le fait que les programmes d'aide périclitaient peu de temps après le départ des donateurs était de plus en plus ouvertement condamné par les chercheurs sur le développement et les ONG internationales de développement[288]. Les évaluateurs durent donc élaborer des méthodes d'évaluation

[284] Conseil Scientifique de l'Évaluation, *L'évaluation en développement,* Rapport annuel sur l'évolution des pratiques d'évaluation des politiques publiques, La Documentation française, 1997.
[285] LAVIGNE-DELVILLE, Philippe, MATHIEU, Marilou, SELLAMNA, Nour-Eddine, *Les enquêtes participatives en débat : ambitions, pratiques, enjeux*, Coédition Gret/Karthala/Icra, 2000, 544p.
[286] *Ibid*.
[287] *Ibid.* pp. 129-142.
[288] CERNEA, Michael. M, *la Dimension humaine dans les projets de développement*, Karthala, 1999, 586p, pp. 26-29.

permettant de juger la pérennité des projets plusieurs années après leur achèvement : les évaluations d'impact. L'évaluation apparût donc, encore une fois, aux décideurs comme un instrument de légitimation face à la critique internationale de leurs politiques.

La Banque mondiale commença à mener des évaluations d'impact dès 1979[289], mais la démarche prit de l'ampleur après une méta évaluation[290] d'impact réalisée en 1985 synthétisant les évaluations déjà réalisées[291]. En effet, cette méta évaluation démontra que 40 % des projets mis en œuvre avec succès quelques années auparavant, avaient périclité sur le long terme. Elle démontra également que dans de nombreux cas, ce n'était pas la population ciblée, le plus souvent les personnes les plus pauvres, qui bénéficiait le plus de l'aide, mais les populations plus favorisées qui étaient dans une meilleure position pour capturer les bénéfices de l'aide au développement. Les conclusions de cette méta évaluation reprenaient celles d'évaluations déjà réalisées durant les décennies précédentes. Dès lors, la Banque mondiale mit en place des groupes de réflexion pour améliorer l'équité des programmes d'aide envers les populations les plus vulnérables. Le retentissement de cette méta évaluation conduisit à la popularisation de la méthode d'évaluation d'impact. En 1993, la Banque mit en place un groupe de travail pour développer une méthodologie normalisée d'évaluation d'impact[292].

Cet exemple montre comment le savoir accumulé par une série d'évaluations peut faire émerger de nouveaux enjeux de développement. En effet, déjà consciente de ce problème du manque de durabilité des programmes d'aide grâce à des évaluations passées et alertée par la société civile (chercheurs et ONG) sur la faible durabilité de ses programmes d'aide, la Banque mondiale réalisa une méta évaluation d'impact qui apporta des preuves « rigoureuses » supplémentaires d'un dysfonctionnement important, auquel il fallait remédier. L'évaluation contribua à prouver la légitimité des enjeux de l'équité et de la durabilité des projets et à en faire un aspect central de la réflexion internationale sur l'aide au développement. Cet usage « mobilisateur » de l'évaluation, peu abordé dans la littérature sur l'évaluation, est en fait courant. Toutefois, il faut souvent, soit une évaluation de grande ampleur et à fort retentissement, soit plusieurs travaux d'évaluation, pour que cet effet soit ressenti.

Les évaluations d'impact constituaient un défi méthodologique pour les évaluateurs. En effet, elles n'ont pas vocation à traiter du processus de mise en œuvre de l'aide et de ses résultats directs, comme le font les évaluations *ex post* classiques. Elles portent sur les effets à long

[289] Voir chapitre précédent.
[290] Pour une définition de ce terme, se référer au lexique en annexe 7.
[291] WORLD BANK, *Annual Report of Project Performance Results*, Report n°5859, OED, 1985.
[292] CRACKNELL, 2000, *op. cit.* p. 239.

terme, dérivés, non prévus, positifs ou négatifs des projets, ainsi que leurs effets distributionnels sur les bénéficiaires directs et indirects.

De plus, la notion de durabilité en elle-même, définie par le CAD-OCDE dans un rapport de 1986, était assez floue : « un projet de développement est durable lorsqu'il est capable d'entraîner un niveau approprié de bénéfices sur une période de temps étendue après que l'assistance majeure financière, managériale et technique du donateur étranger est terminée »[293]. Où se situe « un niveau approprié » ? À quelle temporalité fait référence « une période de temps étendue » ? Que constitue l' « assistance majeure » ? Il était difficile pour les évaluateurs de juger de la durabilité des projets à l'aune de cette définition. L'un d'entre eux a témoigné : « il n'est pas facile de déterminer quel est le bon moment pour mener une évaluation d'impact, et cela varie probablement d'un projet à l'autre. Si elle est menée trop tôt après la fin du projet, alors il est probable que tous les impacts ne se soient pas fait ressentir. Mais si elle est menée trop tardivement, les facteurs de changement non reliés au projet se multiplient et peuvent avoir modifié significativement l'environnement »[294]. Un autre défi méthodologique était de savoir jusqu'à quel cercle de bénéficiaires indirects pousser l'analyse : communautaire, régional ou national ?

Dans un contexte où l'objectif universel donné à l'aide au développement était de réduire la pauvreté, les donateurs attendaient surtout des évaluations d'impact qu'elles donnent des résultats sur les effets distributionnels des projets sur les populations pauvres. Les évaluateurs durent donc décider de la méthode à adopter pour mesurer ces effets. Deux options, théorisées par les évaluateurs de la Banque mondiale, s'offraient à eux :

- l'approche explicative qui visait la quantification systématique des changements analysés. Elle impliquait l'utilisation de méthodes expérimentales[295], notamment la constitution de groupes test (bénéficiaires du projet) et de groupes de contrôle (non-bénéficiaires du projet) permettant d'évaluer les écarts de développement entre ces deux groupes ;
- l'approche exploratoire qui optait pour une technique plus simple impliquant la documentation des changements apparus sans essayer d'établir un contrôle et une quantification. Cette approche était moins coûteuse en temps et en argent que l'approche explicative.

[293] DAC-OECD, « Glossaire des termes utilisés dans l'évaluation », dans *Méthodes et procédures dans l'évaluation de l'aide,* OCDE, 1986.
[294] DAC-OECD, 2000, *op. cit.*
[295] Pour un développement de la méthodologie des évaluations expérimentales voir lexique.

La première option fut rapidement rejetée par les évaluateurs de la Banque mondiale, compte tenu des coûts et du temps nécessaires à la mise en œuvre de cette méthodologie. Selon Susan Stout, ancienne évaluatrice à la Banque mondiale, elle n'était pas réaliste car « *avec un large nombre de projets d'une grande diversité à évaluer et une demande grandissante pour les évaluations d'impact, une méthodologie et une procédure qui permettent, de manière rapide et fiable, des prises de décisions mieux informées est préférable à un petit nombre d'analyses rigoureuses qui échouent souvent à aboutir* »[296]. D'une manière plus radicale, Thomas Cook, chercheur britannique en évaluation des politiques publiques, déclara dans une allocution à la Banque mondiale en 2000 : « *nous pouvons invoquer plusieurs arguments contre l'utilisation des méthodes expérimentales. Le premier est simplement que ce n'est pas faisable* »[297]. Nous observons donc, dans les années 1990, un grand scepticisme de la plupart des évaluateurs par rapport aux évaluations expérimentales dans le domaine de l'aide au développement.

L'approche exploratoire lui fut donc préférée. Pour mener les analyses d'impact, les évaluateurs se fondaient sur des enquêtes auprès des populations les plus pauvres, avant et après les projets, pour essayer d'avoir une vision la plus juste possible de l'amélioration de leurs conditions de vie. Cette méthodologie était très complexe à mettre en œuvre, même si elle ne demandait pas d'instruments statistiques avancés. En effet, les populations pauvres sont généralement les bénéficiaires les moins accessibles de la société et les plus isolés. Parvenir à interroger les populations les plus vulnérables était un défi pour les évaluateurs, qui n'avaient souvent pas de formation en anthropologie. Point faible que les anthropologues du développement s'empressèrent de dénoncer, conduisant à délégitimer l'expertise des évaluateurs. Ces derniers étaient conscients de la limite de leurs compétences en la matière, comme en témoigne, dans ses mémoires, un évaluateur d'USAID : « ces populations subissent une grande pression rien que pour survivre et les enquêteurs de terrain ne sont pas toujours accueillis à bras ouverts. Gagner la confiance des populations n'est pas aisé et le processus ne peut pas être hâtif. Les évaluations d'impact ne sont pas adaptées au format des missions d'évaluation de deux ou trois semaines. Rien qu'atteindre des parties reculées des pays peut prendre un temps considérable. Enfin, les officiels nationaux peuvent être réticents à laisser des agents externes contacter directement les bénéficiaires »[298]. Malgré les difficultés qu'ils rencontraient pour mener ces enquêtes, les évaluateurs étaient très enthousiastes quant aux bénéfices de cette approche pour les populations bénéficiaires : « un des grands avantages de

[296] Entretien avec une évaluatrice retraitée de l'OED, Banque mondiale, 20 novembre 2013.
[297] Thomas Cook, chercheur britannique en évaluation des politiques publiques, *Allocution lors du séminaire de la Banque mondiale célébrant les 30 ans de son département d'évaluation*, archives de la Banque mondiale, 2000.
[298] CRACKNELL, 2000, *op. cit*, pp. 123-124.

l'évaluation d'impact est qu'elle implique généralement les autorités locales et les bénéficiaires dans la mise en œuvre des enquêtes. Non seulement cela permet un transfert de capacités dans la technique des enquêtes statistiques, mais tous les résultats restent dans les pays en développement et sont disponibles pour les recherches futures menées localement »[299]. Or, ces enquêtes, qui se résumaient souvent à la collecte de données quantitatives et à quelques entretiens, ne nous apparaissent pas suffisantes pour permettre une appropriation et un renforcement des capacités des parties prenantes locales, car elles n'étaient pas accompagnées de formations pour permettre une implication active des officiels et des chercheurs locaux dans les processus de collecte et de traitement de données.

Dans les années 1990, les évaluateurs avaient donc toujours une approche qualitative des évaluations d'impact. Cette approche sera radicalement modifiée dans les années 2000 avec, d'une part une demande politique croissante pour les méthodologies expérimentales, et d'autre part un « lobbying » puissant de certains chercheurs en économétrie cherchant à faire prévaloir ces méthodes d'évaluation pour s'imposer dans le marché international grandissant de l'évaluation de l'aide au développement[300].

C. Des standards de l'évaluation influencés par la doctrine du New Public Management

Nous avons choisi de faire débuter ce chapitre en 1991, car cette année marque la publication par le CAD-OCDE du document qui a fondé les « bonnes pratiques » de l'évaluation jusqu'à aujourd'hui. Selon nous, la production de ce référentiel provenait d'une volonté de normaliser la fonction d'évaluation dans les organismes d'aide, afin d'y renforcer le statut et la crédibilité des évaluateurs. Nous allons désormais analyser le processus d'élaboration des critères de cette « bonne pratique » de l'évaluation et la manière dont ils ont été mis en place par les agences d'aide bilatérales.

[299] World Bank, 2000, *op. cit.*
[300] Entretien avec un évaluateur à l'AFD spécialisé dans les méthodes expérimentales, 19 novembre 2013.

- **La formulation de « bonnes pratiques » d'évaluation au niveau multilatéral**

Les *Standards de qualité pour l'évaluation du développement* furent élaborés par le réseau d'évaluateurs *Working Group on Aid Evaluation* du CAD-OCDE. Plusieurs motivations ont incitées ces évaluateurs à mettre en place un ensemble de bonnes pratiques pour normaliser la pratique de l'évaluation entre les agences d'aide. Un membre de ce réseau, raconte qu'« *une des motivations était d'améliorer la qualité des évaluations, afin de renforcer leur crédibilité, alors que le statut des évaluateurs était controversé dans les agences d'aide* »[301]. Une seconde raison était de faciliter la comparaison et le partage des évaluations entre les pays donateurs. En effet, une des missions du groupe de travail était de capitaliser des leçons d'expérience à partir des évaluations menées par chaque donateur. Or, la variété des formats et des contenus des évaluations rendaient ce travail difficile. De plus, la mise en place de standards communs devait faciliter l'émergence d'évaluations conjointes entre les bailleurs[302].

Dès 1985, l'évaluateur représentant le Royaume-Uni proposa l'élaboration d'un guide de critères communs sur les méthodes et les procédures d'évaluation. Cette demande fut rejetée par le Secrétariat du CAD, selon qui la production d'un rapport normatif sur les aspects procéduraux et méthodologiques des évaluations ne rentrait pas dans le mandat du groupe. De plus, certains membres, notamment la France et l'Allemagne, étaient réticents à l'idée de produire un tel document directif. Il fut alors décidé qu'au lieu d'édicter un ensemble de critères communs, le travail du groupe se limiterait à la collecte « d'informations sur les méthodes et les procédures existantes d'évaluation »[303].

Très rapidement cependant, les évaluateurs du réseau se heurtèrent à l'absence de standards communs aux évaluations qui rendait difficile la comparaison des évaluations et la capitalisation de connaissances, notamment sur les nouveaux enjeux du développement. Faute de méthodologies et de résultats comparables, les deux rapports de synthèse sur les évaluations sur le genre et la durabilité qu'entreprit le groupe de travail ne parvinrent pas à apporter de leçons intéressantes[304].

Ces mauvaises expériences conduisirent le groupe à décider, en 1989, de travailler à la conception d'un guide de standards sur l'évaluation du développement, malgré les réserves de

[301] Entretien avec un membre du réseau d'évaluateur du CAD-OCDE, 16 novembre 2012.
[302] DAC-OECD, 2006, *op. cit.*
[303] DAC-OECD (b), 1986, *op. cit.*
[304] DAC-OECD (b), 1986, *op. cit.*

certains membres. Ce travail dura deux ans pour aboutir, en 1991, à la publication du document final : Les principes du CAD sur l'évaluation de l'aide au développement[305].

Néanmoins, compte tenu des réticences de certains membres du réseau à émettre ce document d'harmonisation des pratiques de l'évaluation, les membres du groupe s'étaient accordés sur le fait de ne pas produire un document trop détaillé et prescriptif. Selon un membre de ce réseau, « *le rapport avait plutôt vocation à émettre des principes généraux sur l'évaluation, assez souples pour que chaque donateur puisse se les approprier. L'idée était de ne pas mettre trop de barrières aux agences d'aide. Les membres du groupe pensaient qu'il fallait des normes, mais qu'elles ne devaient pas empêcher l'innovation. C'est pour cette raison que le rapport était succinct et que les standards n'étaient pas très détaillés* »[306].

Le rapport dépassa l'ambition que lui donnèrent ses concepteurs et fut fondamental dans l'histoire de la diffusion d'une « bonne pratique » internationale de l'évaluation dans le secteur de l'aide au développement. Tout d'abord, il proposa la première définition officielle de l'évaluation, toujours utilisée aujourd'hui : « une évaluation est un jugement, aussi systématique et objectif que possible, d'un projet, programme ou politique en cours ou achevé, de son format, sa mise en œuvre et ses résultats. L'objectif est d'en déterminer la pertinence, l'efficience, l'efficacité, l'impact et la durabilité. Une évaluation devrait fournir une information crédible et utile, permettant l'intégration de leçons apprises dans le processus de prise de décision par les bénéficiaires et les donateurs »[307]. Cinq critères ont été tirés de cette définition, qui forment la norme de l'évaluation. Une évaluation de qualité doit traiter de : la pertinence, l'efficacité, l'efficience, l'impact et la durabilité des projets et programmes évalués. Dès lors, les évaluateurs de l'aide au développement s'attachèrent à inclure systématiquement chacun de ces critères dans leurs évaluations, les utilisant comme une protection contre d'éventuelles critiques sur la qualité de leurs travaux. Les avis des évaluateurs sur l'utilité du cadre normatif produit par le CAD-OCDE sont variés. Certains le considèrent comme un « carcan » qui empêche l'innovation et qui oblige les évaluateurs à traiter dans chacun de leurs travaux l'ensemble des principes du CAD (efficacité, utilité, efficience, pertinence), au lieu d'approfondir certaines de ces dimensions seulement, selon l'objet des évaluations[308]. Notre recherche nous a amené à penser que si les bailleurs de fonds ont tendance à utiliser ces critères de manière systématique, sans en comprendre l'utilité et la portée et sans les adapter à leurs propres cultures

[305] *Ibid*.
[306] Entretien avec un membre du réseau d'experts en évaluation du CAD-OCDE, 16 novembre 2012.
[307] DAC-OECD, *Les principes du CAD-OCDE sur l'évaluation de l'aide au développement*, 1991.
[308] Entretien avec un ancien évaluateur au ministère de la coopération, consultant auprès d'organisations internationales, 21 mars 2013.

organisationnelles, c'est parce qu'ils ont voulu les appliquer le plus rapidement possible après leur élaboration pour donner l'image de « bons élèves ». En effet, il y eut peu de dialogue sur ces principes entre les évaluateurs et le personnel des agences d'aide lors de leur adoption. Dès leur ratification, les organismes de développement mirent en place des réformes pour se conformer à ces nouveaux critères. Ainsi, ils ont été intégrés aux procédures des départements d'évaluation, sans être interprétés, appropriés et adaptés par le personnel. Les appels d'offres et les rapports des évaluations, dont la structure suit systématiquement la liste de ces principes, témoignent de la référence systématique aux principes du CAD-OCDE. Ce systématisme a été dénoncé par les concepteurs des principes eux-mêmes : « *il faut faire attention ne pas perdre le sens du projet pendant l'évaluation. Aucun évaluateur n'arrive à bien traiter tous les critères du CAD dans une évaluation, il y a toujours une sélection* »[309]. Les membres du réseau d'évaluateurs du CAD-OCDE tendent à regretter cette lecture des standards : « *les critères du CAD n'ont pas été conçus pour être une camisole. C'est une dérive de leur utilisation que de les appliquer de manière automatique, sans réflexion et sans adaptation. Les bailleurs de fond les prennent comme des rituels et font preuve d'un manque d'imagination qui interpelle le CAD* »[310].

D'autres évaluateurs, souvent externes aux organismes de développement, estiment que ce cadre est assez ouvert pour permettre à l'évaluateur de garder sa liberté d'analyse[311]. Ainsi, si les bonnes pratiques du CAD sont considérées comme utiles par la majorité des évaluateurs, ces derniers revendiquent la possibilité d'en moduler l'application selon les besoins de chaque évaluation, ce que semble avoir du mal à accepter la hiérarchie des agences d'aide souhaitant que les évaluations soient irréprochables en termes de rigueur méthodologique et soient les plus exhaustives possibles.

En plus de normaliser le champ des aspects traités par les évaluations, les principes du CAD établissent des bonnes pratiques pour guider les procédures d'évaluation :

- les agences d'aide doivent avoir une politique d'évaluation qui décrit clairement le statut, le rôle et les méthodes des évaluations, ainsi que les responsabilités des évaluateurs ;
- les processus d'évaluation doivent être indépendants de la prise de décision, de l'allocation et de la gestion de l'aide au développement ;

[309] Entretien avec un membre du réseau d'évaluateur du CAD-OCDE, 16 novembre 2012.
[310] Entretien menés auprès du CAD-OCDE, voir liste des entretiens en annexe 8.
[311] Entretien avec un directeur de cabinet d'évaluation des politiques publiques basé en France, le 9 décembre 2014.

- les processus d'évaluation doivent être le plus ouverts possible et leurs résultats doivent être rendus publics ;
- les évaluations doivent être accompagnées de procédures de suivi pour assurer leur utilisation ;
- les évaluations doivent être, dès que possible, réalisées en partenariat avec les bénéficiaires de l'aide et/ou les donateurs partenaires[312].

Ces standards et bonnes pratiques devaient permettre aux donateurs de contrôler et de juger de la fiabilité des évaluations qu'elles externalisaient à des consultants privés.

Les membres du réseau d'évaluateurs du CAD-OCDE considèrent aujourd'hui les Standards de qualités sur l'évaluation de l'aide au développement comme « les critères les plus célèbres et les plus largement adoptés de l'aide au développement. Ils sont utilisés par la majorité des organisations bilatérales et multilatérales, ainsi que les organisations non-gouvernementales »[313]. Nos discussions avec des acteurs de l'aide qui ne sont pas des évaluateurs nous ont fait relativiser cette affirmation. Les critères du CAD-OCDE sont souvent vaguement connus, voire méconnus, par les opérateurs de l'aide.

En 1998, le CAD-OCDE effectua une revue transversale pour juger de l'application des standards de 1991 par les agences d'aide de ses pays membres[314]. Les représentants du CAD présentèrent positivement les résultats de cette étude : « ce bilan a permis de constater que de nombreux membres du réseau ont restructuré leurs services centraux d'évaluation afin de faire évoluer leur mission et davantage cibler leur action sur l'efficacité de l'aide (…). Les membres ont souligné l'intérêt des principes et leur pertinence quant à l'orientation de leurs travaux, voire, dans certains cas, la réorganisation de leurs services d'évaluation (…). La révision des principes a permis de constater que la majorité des membres s'y conforment dans une large mesure »[315].

Cependant, nous tirons de notre lecture de cette revue des conclusions bien plus négatives. Tout d'abord, il nous apparaît paradoxal de conclure que les membres « soulignent l'intérêt des principes » alors que la revue dévoile plus loin que « les membres déclarent qu'ils ne sont pas familiers avec les principes du CAD pour l'évaluation. Dans quelques cas seulement, les personnels interrogés avaient lu les principes du CAD »[316]. Cette remarque

[312] DAC-OECD, 1991, *op. cit.*
[313] DAC-OECD, 2013, *op. cit.*
[314] DAC-OECD, 1998, *op. cit.*
[315] *Ibid.*
[316] *Ibid.*

renforce notre constat selon lequel les agences d'aide ont intégré les principes du CAD et réformé en fonction d'eux leurs départements d'évaluation sans qu'ils soient pour autant compris et appropriés par leurs utilisateurs.

De manière générale, si les bailleurs de fonds affichaient leur volonté de modifier leur approche de l'évaluation pour se conformer au mieux aux principes du CAD, la mise en œuvre de cette volonté était limitée. En ce qui concerne les principes d'impartialité et d'indépendance de l'évaluation par exemple, les évaluateurs internes aux agences d'aide interrogés dans la revue reconnaissaient « négocier les conclusions et recommandations » et tout faire pour « l'obtention d'un consensus », ce qui nuit aux principes du CAD[317].

Notre analyse de la revue fait également ressortir que les principes de crédibilité et d'inclusion de l'évaluation étaient peu appliqués. En effet, ces principes impliquent la participation des pays bénéficiaires qui restait rare dans les années 1990 : « alors que la participation des bénéficiaires est considérée comme essentielle à la crédibilité des évaluations des donateurs, elle est rarement mise en pratique. En conséquence, la revue estime que le respect de ce principe par les membres du CAD-OCDE est insuffisant »[318].

Concernant l'utilité des évaluations, hormis les évaluations de projets qui avaient une utilité opérationnelle directe, les évaluations de programmes, sectorielles, ou par pays étaient peu utilisées. Les explications données par le personnel des agences d'aide pour justifier cette faible utilisation des évaluations étaient variées : « les rapports sont trop volumineux et illisibles », « les résumés ne sont pas clairs et précis » ; « l'accès à l'information est trop coûteux en temps »[319]. La revue concluait que, dans un nombre significatif d'agences d'aide, l'influence des évaluations était minimale.

Enfin, la coordination entre donateurs, devant mener à une meilleure pertinence des évaluations (en évitant que des programmes « multi-bailleurs » soient évalués plusieurs fois), nous apparaît être le principe le plus faiblement appliqué. En effet, nous trouvons dans le texte de la revue des formules du type : « même s'il existe quelques exemples de collaborations réussies, les donateurs sont généralement peu enthousiastes à réaliser des évaluations conjointes »[320].

---

[317] *Ibid*.
[318] DAC-OECD, Working Group on Evaluation, 1998, *op. cit.*
[319] *Ibid.*
[320] *Ibid*.

Nous constatons ainsi un large écart entre le discours du CAD-OCDE sur la mise en œuvre des « bonnes pratiques » de l'évaluation et la réalité. Nous ne remettons pas en cause la volonté qu'avaient des dirigeants des agences d'aide de se conformer aux standards du CAD-OCDE. En témoigne les références multiples qui y étaient faites dans les documents stratégiques des politiques de coopération pour le développement des pays membres. Toutefois, dans les années 1990, cette volonté eut des difficultés à dépasser le stade de l'intention, ce qui montre que la diffusion d'une culture de l'évaluation demande un temps long.

Nous allons désormais démontrer que la principale raison de cette faible mise en pratique des standards du CAD provenait du contexte des politiques publiques de la plupart des États occidentaux, où l'évaluation n'était pas instituée dans la culture administrative. Un premier indice qui appuie cette idée est que les bailleurs de fonds les plus convaincus de l'utilité des principes d'évaluation du CAD et qui les appliquaient avec le plus de rigueur étaient le Royaume-Uni, les États-Unis, le Danemark, ou encore la Suède, où la culture du New Public Management et de l'évaluation était bien ancrée. Cette influence du New Public Management conduisit à focaliser la fonction de l'évaluation sur la mesure de la performance et le contrôle de la mise en œuvre des programmes d'aide.

- **L'influence du New Public Management sur l'évaluation : les cas d'étude de la France et du Royaume-Uni**

Afin de comprendre en quoi la diffusion internationale des pratiques du New Public Management[321] eut une influence sur la diffusion d'une culture normalisée de l'évaluation dans les organismes d'aide au développement à travers le monde, il convient de revenir sur le fond de ce concept et l'idéologie qui l'entoure.

Le New Public Management est un modèle de gestion des politiques publiques qui a pour principe fondateur la négation d'une distinction entre les méthodes de gestion des organismes publics et privés[322]. Selon ce modèle, les organismes publics doivent adopter des procédures routinières de gestion, organisées autour d'une mesure régulière de la performance. Cette croyance, selon laquelle la décision rationnelle nécessite une gestion par la mesure et le contrôle, remonte au XIXème siècle, lorsque les premières théories du « gouvernement par les

[321] Pour une analyse plus approfondie des principes du New Public Management, voir : FERLIE, Ewan, *The New Public Management in Action*, Oxford University Press, 1996, 287p ; BARZELAY, Michael, *The New Public Management: Improving research and policy dialogue*, Russel Sage Foundation, 2001, 218p.
[322] OSBORNE, David, GAEBLER, Ted, *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector*, Plume, 1993, 432p.

nombres » furent développées, afin de réguler les comportements des sociétés, par exemple dans le secteur de la santé publique[323].

Nous constatons, dans les années 1990, un regain de popularité de ce modèle de gestion des politiques publiques qui se généralisa dans les États occidentaux, après avoir été adopté au Royaume-Uni sous le gouvernement Thatcher et aux États-Unis sous le gouvernement Reagan[324]. Ce succès du New Public Management tenait, selon nous, du contexte de crise mondialisée des années 1990, où les incertitudes sur l'efficacité des politiques engagées étaient grandes, et où les gouvernements voulaient réduire les dépenses publiques et resserrer leur contrôle sur la gestion et les résultats des politiques publiques. Nous identifions clairement le New Public Management comme un modèle gestionnaire « de crise ». Ulrich Beck développe l'idée selon laquelle les sociétés ont tendance à devenir plus prudentes et réticentes face à la prise de risque, avec des individus et des organisations qui développent une ingénierie statistiques, afin d'éviter que des mauvais résultats ne leur soient attribués[325].

Les principes technocratiques du New Public Management [326] entraînèrent une refonte de la structure organisationnelle des organismes publics qui eut une influence significative sur l'évolution de l'évaluation des politiques publiques.

Le modèle du New Public Management préconise que la baisse des coûts de la production de services publics peut se faire sans diminuer (et même en augmentant) la qualité de ces services. Il met en avant l'importance du citoyen bénéficiaire qui est considéré comme un « client » ayant le droit de porter un jugement sur les services qui lui sont fournis. La mise en œuvre des politiques doit être précédée et suivie de consultations avec les citoyens pour connaître leurs besoins *a priori* et leur satisfaction *a posteriori*. Nous constatons que les procédures participatives de gestion des politiques publiques se généralisèrent et s'institutionnalisèrent durant les années 1990[327]. C'est ainsi que fleurirent des manuels

[323] FOUCAULT, Michel, *Du Gouvernement des vivants*, Cours au Collège de France (1979-1980), Gallimard-Seuil, 2004, 380p.
[324] EYBEN, Rosalind, *Uncovering the politics of evidence and results, a framing paper for development practitioners*, Big Push Forward, avril 2013.
[325] BECK, Ulrich, *La société du risque : sur la voie d'une autre modernité,* Flammarion, 2008, 521p.
[326] Les principes fondamentaux du New Public Management sont : i) une mise en œuvre segmentée des politiques publiques sous la forme de missions déléguées à des unités autonomes de management ; ii)la fragmentation des responsabilités et une exigence généralisée de redevabilité, à la fois verticale vers les autorités responsables, horizontale vers les entités de contrôle des organismes publics, et sociale envers les utilisateurs des politiques publiques ;iii)la création d'incitations pour promouvoir l'efficacité individuelle ; iv) la définition d'objectifs et la systématisation des activités de mesure de la performance fondées sur des ensembles d'indicateurs qui rendent opérationnels les objectifs fixés en détaillant des bonnes pratiques pour les atteindre ; v) une mise en œuvre des politiques publiques guidée par une exigence double : la qualité la plus haute et les coûts de « production » les plus bas.
: HOOD, Christopher, "A public management for all seasons", *Public Administration*, vol. 69, spring 1991, pp. 3-19; FERLIE, Ewan, *The New Public Management in Action*, Oxford University Press, 1996, 287p.
[327] HIBOU, Béatrice, 2012, *op. cit.*

formulant des cadres de procédures, de directives, de méthodes « scientifiques » et d'indicateurs pour mener les consultations populaires[328]. Nous soulignons ici l'image technicisée et formalisée qui fut donc donnée à l'acte fondamentalement politique qu'était la consultation entre les dirigeants et les citoyens. Selon nous, cette image a été construite avec la volonté de faire apparaître ces consultations comme des processus encadrés par une technique scientifique et donc de les rendre plus crédibles.

La vision du New Public Management de l'efficacité des politiques publiques nous apparaît limitée, dans la mesure où elle réduit les facteurs de réussite d'une politique à des considérations techniques en faisant abstraction, d'une part des facteurs internes relevant du politique ou de la psychologie des acteurs, et d'autre part des facteurs contextuels externes pouvant influencer la capacité des organismes à atteindre leurs objectifs. Pourtant, ce modèle de gestion rationnelle et « non-bureaucratique » des administrations eut beaucoup de succès dans les années 1990, dans tous les secteurs des politiques publiques qui cherchaient à réduire leurs dépenses.

Avec le développement du New Public Management, les organisations publiques durent se doter de mécanismes de contrôle et d'évaluation internes de leurs activités pour prouver leur performance de manière transparente. En plus de ces mécanismes internes, les agences externes chargées de classer, de contrôler et d'évaluer les organisations publiques se multiplièrent[329]. Cette hausse du contrôle et de l'évaluation, interne et externe, eut plusieurs conséquences sur la gestion des politiques publiques théorisées par Michael Power[330] :

- l'adoption de schémas standardisés de management du risque devint une obligation pour que les organisations soient considérées comme légitimes ;
- les organismes publics devinrent de plus en plus attentifs au « risque de réputation » qu'entraîne le choix des projets et programmes mis en œuvre ;
- les décisionnaires des organisations se placèrent dans une logique performative de court terme.

Nous estimons que ce modèle de gestion des politiques publiques généra une forte demande d'évaluations de la part des organismes publics, afin de prouver leur performance, de

[328] *Ibid*.
[329] Dans le secteur de l'aide au développement, ces pratiques de contrôle et de classement externe des agences d'aide au développement se développent de manière significative tardivement, dans les années 2000. Ce phénomène sera donc analysé dans le chapitre suivant.
[330] POWER, Michael, *Organized Uncertainty, Designing a world of risk Management*, Oxford University Press, 2007, 248p.

rationaliser leurs modes de fonctionnement et donc leur efficience et de limiter les risques de leurs activités dans un contexte international perçu comme instable.

Nous allons désormais étudier comment la diffusion des principes du New Public Management influença la pratique de l'évaluation des politiques publiques au Royaume-Uni et en France.

Le Royaume-Uni est le pays de naissance du New Public Management et dans lequel il a eu la plus grande influence sur la gestion des politiques publiques. Théorisé dans les années 1970, il fut mis en pratique à partir des années 1980 dans un contexte de crise budgétaire de l'État[331]. Le gouvernement de Margaret Thatcher s'engagea à privatiser de nombreux organismes publics, afin de « limiter les gaspillages et augmenter la qualité des services à la population ». Plusieurs programmes politiques furent mis en place suivant cette logique d'économies budgétaires, de rationalisation et de hausse de la transparence de la fonction publique. En 1988, le programme politique *Next Steps* créa de nombreuses agences externes à l'État chargées de missions de services publics. Leur fonctionnement était proche de celui des entreprises et elles étaient censées fournir aux citoyens des services de meilleure qualité pour un coût moindre. Cette approche néolibérale de la gestion des politiques publiques se poursuivit dans les années 1990 sous le gouvernement de John Major. Des documents de réforme du secteur public furent produits pour institutionnaliser la pratique du New Public Management : la *Citizen's Charter* de 1991 plaçait le citoyen au centre de l'évaluation de l'administration ; le *Deregulation and Contracting Out Act* de 1994 visait à réduire les coûts et la bureaucratisation de l'administration ; ou encore le *Modernising Government White Paper* de 1999, encourageait une fonction publique efficace, efficiente mais aussi équitable. Dans ces documents réformateurs, nous retrouvons tous les principes du New Public Management : dérégulation, place centrale donnée à la satisfaction du citoyen, équilibre entre la qualité et l'efficience des politiques publiques et réductions budgétaires.

Parallèlement à l'adoption de ces réformes, la fonction de l'évaluation se systématisa au Royaume-Uni. De nombreux chercheurs en politique publique font le constat d'un effort accru, dans les années 1990, pour la détermination d'objectifs, le suivi des performances et l'évaluation régulière des politiques publiques[332]. La politique du New Public Management servait les intérêts des évaluateurs. Elle mit en avant les principes de mesure de la performance,

[331] AMAR, Anne, BERTHIER, Ludovic, *Le nouveau Management Public : Avantages et limites*, CEROG, 2006, 13p.
[332] Voir par exemple CARTER, Nail, KLEIN, Patricia, DAY, Rudolf, *How Organizations Measure Success: the use of performance indicators in Government*, Routledge, 1992, 208p; POLLIT, Summa, "Trajectories of Reform: Public Management Change in Four Countries", *Public Money and Management*, vol. 17, n°1, janvier-mars, pp. 1-13.

de contrôle et d'une meilleure transparence envers les citoyens, autant d'enjeux que sert la fonction d'évaluation.

Nous allons démontrer que la principale influence du New Public Management sur l'évaluation des politiques publiques est qu'elle lui donna un caractère managérial. En effet, dans ce cadre, l'évaluation ne servait plus uniquement à la mesure ponctuelle de résultats de programmes de politiques innovants ou risqués, elle devenait une activité routinière de la gestion des administrations publiques qui étaient contraintes d'évaluer en continu leur fonctionnement interne et leur performance. De plus, il était commun que les évaluations recommandent elles-mêmes une systématisation de la pratique de l'évaluation au sein des organisations, ce qui entraîna une logique d'auto-renforcement de la fonction d'évaluation. En conséquence, l'évaluation « se répandit dans l'ensemble des administrations britanniques »[333].

Si le New Public Management eut une influence indéniable sur l'accélération de la diffusion d'une culture de l'évaluation dans les années 1990 dans les pays européens, la vitesse et les circuits de cette diffusion, ainsi que la forme que prit la fonction d'évaluation dans les organisations publiques varièrent selon les États. Patrice Durand, Éric Monnier et Andy Smith, spécialistes de l'évaluation des politiques publiques en France, exprimèrent bien cette idée : « en France (contrairement au Royaume-Uni), l'évaluation a été institutionnalisée selon des modalités qui sont typiques du style d'administration publique de ce pays. L'institutionnalisation de l'évaluation a précédé son application à des politiques concrètes »[334].

En France, le début des années 1990 fut, comme au Royaume-Uni, marqué par une réforme majeure des politiques publiques qui allait vers une rationalisation des pratiques et une promotion de l'évaluation. En février 1989, le premier ministre Michel Rocard lança une politique de « renouveau du service public ». Elle s'organisait autour de quatre axes : une politique de relations du travail rénovée, une politique de développement des responsabilités, un devoir d'évaluation des politiques publiques, une politique d'accueil et de service à l'égard des usagers[335]. Même si l'influence du New Public Management était moins claire et directe qu'au Royaume-Uni, nous retrouvons bien dans cette réforme les principaux critères de ce modèle, dont une mise en avant du rôle de l'évaluation. Dans le cadre de cette réforme, le 22 janvier 1990, fut publié un décret relatif à l'évaluation des politiques publiques. Il mit en place un dispositif d'évaluation interministériel ayant vocation à diffuser la culture et la pratique de

[333] HOOD, 1992, *art. cit.*
[334] DURAN, Patrice, MONNIER, Eric, SMITH Andy, « Evaluation à la française, Toward a New Relationship between Social Science and Public Action », *Evaluation*, vol. 1, n°1, juillet 1995, pp. 45-63.
[335] Premier ministre, *Circulaire du 23 février 1989 relative au renouveau du service public*, 23 février 1989.

l'évaluation dans les différents ministères. Ce dispositif s'appuyait sur un Conseil scientifique de l'évaluation[336] chargé de favoriser le développement des méthodes d'évaluation et de définir une déontologie en la matière. Il veillait à la qualité et à l'objectivité des travaux financés par le *Fonds national de développement de l'évaluation*[337]. Malgré la mise en place de ce cadre institutionnel favorable, l'évaluation eut des difficultés à rentrer dans les pratiques des administrations françaises. En 1997, le Conseil scientifique de l'évaluation[338] déclara en avant-propos de son bilan annuel : « la France ne figure plus, à la fin des années 1990, dans le peloton de tête des pays consommateurs d'évaluations. Tout se passe comme si l'élan des années 1980 (…) s'était estompé. Pendant ce temps-là, des pays tels que la Grande-Bretagne, l'Italie ou les Pays-Bas, sans parler de l'Union européenne, connaissent une diffusion croissante de ces usages »[339].

Cette difficulté à diffuser la pratique évaluative en France pourrait relever de deux caractéristiques de la culture de l'évaluation dans ce pays :

- une vision étroite et technocratique de l'évaluation. Selon le décret de 1990, l'évaluation concernerait surtout les « grandes politiques publiques ». Elle était conçue comme un instrument au service des plus hauts décideurs politiques ;
- une démarche très centralisée de l'évaluation. Les commandes d'évaluation devaient venir du plus haut niveau de responsabilité gouvernementale et politique (ministères, inspections générales, etc.).

Une première conséquence fut que, contrairement au Royaume-Uni, l'évaluation ne s'était pas banalisée comme un instrument parmi d'autres pour accompagner la gestion quotidienne des services publics. L'évaluation échouait à se diffuser à l'intérieur des services, elle était « l'attribut du souverain sans s'incorporer à la boîte à outils des fonctionnaires »[340]. Puisque l'évaluation était l'outil du souverain (le gouvernement), si ce dernier se désintéressait de cet outil, le dispositif se bloquait entièrement. Une évaluation commanditée et menée sans la légitimation et l'impulsion des plus hauts niveaux de l'État n'était pas jugée crédible. En conséquence, l'évaluation était considérée comme une mode qui réapparaissait selon le bon vouloir du « souverain », et non pas comme une pratique ancrée dans la routine bureaucratique et indispensable au bon fonctionnement des institutions. C'est pourquoi l'opinion commune

[336] Pour une analyse plus approfondie du rôle du Conseil scientifique de l'évaluation, voir : MATYJASIK, Nicolas, *op. cit.*
[337] *Décret n° 90-82 du 22 janvier 1990 relatif à l'évaluation des politiques publiques*, Article 8.
[338] Il sera remplacé par le Conseil national de l'évaluation en 1998.
[339] Conseil scientifique de l'évaluation, *L'évaluation en développement,* Rapport annuel sur l'évolution des pratiques d'évaluation des politiques publiques, La Documentation française, 1997
[340] *Ibid.*

rapprocha souvent l'impulsion donnée à l'évaluation à la fin des années 1980 à l'arrivée au pouvoir d'un gouvernement socialiste, dont cela aurait été la marotte, plutôt qu'à une évolution durable de la gestion des politiques publiques.

Dans une allocution lors d'une rencontre entre évaluateurs français et britanniques, organisée par le Conseil scientifique de l'évaluation en 1998, Christopher Pollitt, chercheur au *Centre for the Evaluation of Public Policy and Practice* de l'Université Brunel, élabora une typologie des différentes caractéristiques de l'institutionnalisation de l'évaluation en France, au Royaume-Uni et aux États-Unis, dont nous partageons les conclusions, afin d'en déduire des divergences de pratiques et d'utilisation des évaluations[341] :

- aux États-Unis, la culture de l'évaluation s'est diffusée suivant un modèle de marché. L'évaluation est considérée comme un service commercial. Très rapidement, des cabinets d'experts et des chercheurs universitaires se sont saisis de l'enjeu de l'évaluation. Ils entrèrent en compétition pour obtenir les marchés en évaluation externalisée provenant du gouvernement et des organismes paraétatiques ;
- en France, l'évaluation s'est diffusée de manière hiérarchique, du sommet de l'administration vers le bas. L'évaluation est vue comme une fonction très théorique, imposée par les hautes sphères de l'État et peu adaptée aux besoins des praticiens des politiques publiques. Ce sentiment est dû au fait que son institutionnalisation a précédé son application systématique à des politiques concrètes. La création d'institutions chargées de l'évaluation est le point de départ d'une diffusion d'une culture de l'évaluation ;
- au Royaume-Uni, la culture de l'évaluation s'est développée sous la forme de réseaux sectoriels de professionnels. Ainsi, certains secteurs des politiques publiques avaient une culture de l'évaluation bien ancrée dans les années 1990, alors que d'autres étaient plus en retard. La pratique de l'évaluation était donc assez fragmentée, même si des échanges d'expériences intersectoriels se développaient.

Après avoir analysé la tendance généralisée dans les principaux pays de l'OCDE vers l'adoption de principes de gestion des politiques publiques se rapprochant du modèle du New Public Management, nous allons nous intéresser aux effets de cette tendance sur les organismes d'aide au développement, et plus spécifiquement sur leur pratique de l'évaluation.

[341] Conseil Scientifique de l'évaluation, 1997, *op. cit.*

- **La diffusion des « bonnes pratiques » d'évaluation au niveau des agences d'aide bilatérales**

Comme le constatait Jean-David Naudet dans un article publié en février 2012, aucun travail de recherche n'est consacré à l'analyse de l'influence de la normalisation des pratiques de New Public Management dans les démocraties occidentales sur l'évolution des méthodes de gestion de l'aide au développement dans les années 1990[342].

Au début des années 1990, un débat eut lieu au sein du CAD entre les donateurs anglo-américains qui prônaient une approche de l'évaluation inspirée par le New Public Management, où l'évaluation est un outil gestionnaire de mesure des résultats quantitatifs à court terme, et les donateurs français et allemands qui étaient plus favorables à des approches « sociales » et participatives de l'évaluation, à des fins de réflexion partagée sur les intrications complexes des problèmes de développement[343]. Dès la fin des années 1990, la vision anglo-saxonne prévalut, à mesure que les bailleurs de fonds firent face à une demande accrue de résultats de la part de leurs tutelles financières dans un contexte international où le courant de pensée du New Public Management montait en puissance.

Dans ce contexte de montée du New Public Management, des chercheurs en politiques publiques développèrent la « théorie du changement » qui connut un grand succès parmi les évaluateurs de l'aide au développement. La théorie du changement est une méthode de planification des politiques publiques qui consiste à expliquer *a priori* le processus de changement induit par une politique publique en identifiant les liens de causalité entre les résultats à court, moyen et long terme de cette politique. Ces liens doivent être explicités par les parties prenantes de la politique publique en amont de sa mise en œuvre. Ainsi, la théorie du changement requiert des parties prenantes d'une politique publique qu'elles modélisent les résultats attendus avant de décider des modalités d'intervention à prendre pour atteindre ces résultats. L'objectif est de faire prendre des décisions mieux informées aux praticiens des politiques publiques. Selon cette théorie, une fois que les données issues du suivi-évaluation de la politique sont disponibles, les parties prenantes peuvent affiner à nouveau leur théorie du changement à partir des « preuves » apportées. La théorie du changement veut que la planification et le suivi-évaluation soient aussi inclusifs que possible pour prendre en compte les perspectives de toutes les parties prenantes de la politique publique.

[342] *Ibid*.
[343] Entretien avec un évaluateur de l'AFD, spécialisé dans les méthodes expérimentales, 19 novembre 2013.

L'objectif de cette théorie est de rationaliser les politiques publiques en fondant la prise de décision sur l'anticipation de chaînes de causalité et de résultats réalistes. Le suivi-évaluation devient ainsi un instrument central de la mise en œuvre des politiques publiques[344].

Un premier signe de l'introduction de l'idéologie du New Public Management dans le secteur de l'aide au développement dans les années 1990 est que le néologisme « redevabilité » provenant du New Public Management, s'imposa comme un terme courant du langage de l'aide au développement. Ce furent les organismes multilatéraux qui adoptèrent les premiers ce vocabulaire pour donner une image de « gestion saine » de l'aide au développement alors en crise.

Nous avons vu ci-dessus que le New Public Management est un modèle de gestion des politiques publiques qui se prête idéalement aux périodes de crise. Or, dans les années 1990, l'aide au développement connut une forte perte de légitimité. En outre, la mise en œuvre de l'aide qui était segmentée par projets et par pays se prêtait bien à la décentralisation. Enfin, les acteurs de l'aide au développement se multiplièrent et se spécialisèrent : ONG, fondations, fonds verticaux, etc. Une concurrence émergea entre ces acteurs et, en conséquence, ils durent apporter plus de preuves de leur efficacité.

Un deuxième changement important apporté à l'aide au développement par le New Public Management fut l'accent mis sur les bénéficiaires des projets et programmes. Comme nous l'avons vu au début de ce chapitre, l'aide au développement « s'individualisa » et les programmes d'aide se spécialisèrent selon des publics de plus en plus précis : les « pauvres », les « exclus », les « femmes », etc. Dans les années 1990, l'idée commença à naître que les bénéficiaires du Sud avaient le droit d'évaluer la qualité de l'aide qui leur était délivrée. Cette idée fut traduite en pratique par une hausse des évaluations participatives qui se fondaient en grande partie sur les récits des populations bénéficiaires de l'aide et sur l'apprentissage mutuel entre les donateurs et les bénéficiaires[345]. Selon nous, il était toutefois difficile de transposer des pratiques de consultations citoyennes à l'aide au développement tant que les relations entre donateurs et bénéficiaires n'étaient pas régulées par un principe démocratique et que la culture des politiques publiques des pays concernés était éloignée de ces principes participatifs.

---

[344] *Pour* advantage d'informations sur la théorie du changement, voir par exemple : WEISS, Carol, "Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families", in CONNELL, James *et al*. (Eds.) *New Approaches to Evaluating Community Initiatives*, Aspen Institute, 1995, 255p; TAPLIN, Dana, *et al*. *Theory of Change, a Serie of Papers to Support Development of Theories of Change based on Practice in the Field*, ActKnowledge, avril 2013.

[345] JACKSON, Edward T, KASSAM, Yusuf, *Knowledge Shared: Participatory Evaluation in Development Cooperation*, IDRC, Kumarian Press, 1998, 252p.

À la lecture des examens par les pairs du CAD-OCDE, nous constatons que cet organisme exerçait, dans les années 1990, une pression normative forte sur les bailleurs de fonds pour qu'ils adoptent des pratiques de gestion de l'aide répondant aux principes néolibéraux du New Public Management. Les recommandations du CAD-OCDE apparaissent, en effet, clairement inspirées du New Public Management[346] :

- décentraliser la gestion des projets et programmes ;
- responsabiliser le personnel des agences d'aide quels que soient le statut et la position hiérarchique des individus ;
- renforcer la redevabilité interne et externe des organismes d'aide ;
- valoriser l'efficacité individuelle et la prise d'initiative pour sortir d'un mode de gestion « bureaucratique » ;
- mettre en place des « systèmes de mesure et de gestion de la performance » qui matérialiseraient les processus de redevabilité entre les différentes unités des organisations.

Nous pensons ainsi qu'un phénomène de « bureaucratisation néolibérale » gagna la sphère du développement international, caractérisé par : « la montée en puissance de la rationalité technique, l'envahissement des normes du marché et de l'entreprise, la formalisation du gouvernement à distance, l'intensification des opérations d'abstraction (…) et la suprématie d'un système d'évaluation quantitatif »[347]. En effet, la théorie du changement, très liée au New Public Management, eut un grand succès dans les organismes d'aide au développement anglo-saxons (USAID, DFID, Banque mondiale) qui la promurent au sein du CAD-OCDE[348].

Les nouvelles orientations prônées par le CAD-OCDE, que ce soit la décentralisation, la responsabilisation individuelle ou l'amélioration de la transparence, nécessitaient une systématisation des procédures de supervision non-coercitives, incarnées par l'évaluation. C'est ainsi que, sous l'influence de la théorie du changement et du New Public Management, le nombre des évaluations augmenta significativement durant la décennie 1990. D'après les données du CAD-OCDE[349], les unités d'évaluation des organisations de développement menèrent environ 9 300 évaluations dans la décennie[350].

[346] DAC-OECD, Working Party on Evaluation, *Review of DAC principles for development assistance*, 1998.
[347] HIBOU, Béatrice, 2012, *op. cit.* p. 33.
[348] STEIN, Danielle, VALTERS, Craig, *Understanding Theory of Change in International Development*, the Justice and Security Research Programme, London School of Economics, août 2012.
[349] DAC Working Group on Evaluation, 1998, *op. cit.*
[350] On n'inclue pas dans ce nombre les évaluations qui peuvent être menées par les autres divisions et les évaluations menées par des organisations non gouvernementales.

Une autre illustration de la culture du New Public Management peut être trouvée dans une réforme réalisée en France intégrant le principe de de responsabilisation des différents niveaux hiérarchiques des organismes d'aide. Dans les années 1990, la mise en œuvre de la politique d'aide au développement était toujours fragmentée et peu coordonnée entre les différents organismes qui en étaient chargés[351]. Jusqu'alors la CFD était le principal opérateur technique de l'aide sans pour autant avoir de responsabilité décisionnaire sur les projets et programmes qu'elle mettait en œuvre. À la suite de recommandations d'un examen par les pairs du CAD[352] mené par l'Italie et la Norvège et pour se conformer aux travaux européens sur l'efficacité de l'aide, le gouvernement Jospin entreprit de donner plus de responsabilités à la CFD. En 1998, cette dernière devint l'Agence française de développement (AFD), qui sera désormais l'organisme pivot de l'aide au développement, alors que le ministère de la coopération était intégré dans le ministère des affaires étrangères. Le fait que cette réforme ait en partie été motivée par les recommandations d'une revue par les pairs du CAD-OCDE tend à confirmer que l'organisme avait une influence normative accrue. Nous voyons également dans la création de l'AFD un signe de la montée en puissance du modèle du New Public Management en France dans les années 1990.

En effet, le New Public Management était porteur d'une « agencification » de l'État qui s'étendit à toute l'Europe, c'est-à-dire « une multiplication d'agences, entités hybrides à la fois publiques et privées, disposant d'une certaine autonomie envers leur tutelle étatique et d'une légitimité liée à l'expertise technique de leur personnel. Ces agences ont été créées pour surmonter les rigidités des bureaucraties nationales et assurer la mise en œuvre efficace d'une politique dans son environnement »[353]. Cette citation décrit aussi bien les cas du Royaume-Uni, où le DFID fut institué en tant qu'agence gouvernementale séparée du *Foreign and Commonwealth Office*, que de la France, où l'AFD a profité de la disparition du ministère de la coopération pour s'affirmer comme l'institution techniquement la plus compétente en matière d'aide au développement. Pour d'aucuns, c'est grâce à l'action d'Antoine Pouillieute que l'AFD put se positionner comme l'acteur central de l'aide au développement en France[354]. Selon Corinne Balleix, « l'AFD est un pur produit de la bureaucratisation néolibérale du New Public Management qui transforma les fondements de l'APD française. L'AFD est imprégnée, à l'instar d'autres agences européennes, d'un référentiel libéral d'efficacité et de sélectivité de

[351] Le secrétariat d'État à la Coopération, le ministère des Affaires étrangères, la Caisse française de développement, la cellule Afrique de l'Élysée et un service spécialisé à Matignon.
[352] DAC-OECD, 2000, *op. cit.*
[353] BALLEIX, Corinne, « La politique française de coopération au développement : Cinquante ans d'histoire au miroir de l'Europe », *Afrique Contemporaine*, vol. 4, n°236, 2010, pp. 95-107.
[354] Entretien avec un directeur de département de l'AFD, 5 décembre 2013.

l'aide, elle promeut une coopération mêlant prêts et dons, élargie aux biens publics mondiaux et aux pays émergents, dépassant dès lors les critères historiques de l'aide française »[355].

De plus, le fonctionnement de l'AFD était décentralisé : « les tâches des agences locales s'étendirent dans les années 1990 en raison du désir d'un plus grand formalisme, de la mise en place de procédures de contrôle et d'évaluation, afin de fournir de la redevabilité et d'améliorer l'efficacité de l'aide. Le résultat de cette tendance fut un renforcement de la décentralisation des responsabilités et de la prise de décision » [356]. Dans cette citation, nous voyons ressortir les principes de redevabilité, d'efficacité et de responsabilité, si caractéristiques au New Public Management. Nos recherches nous ont amené à déduire que cette structuration de l'AFD fut grandement inspirée par le CAD-OCDE et l'Union européenne qui cherchaient à diffuser le modèle de gestion du New Public Management dans le secteur de l'aide au développement. En effet, le rôle l'Union européenne en matière de normalisation des pratiques d'évaluation des États membres grandit dans les années 1990[357].

Ainsi, nous observons que l'APD française était de plus en plus influencée par la pression normative des organisations multilatérales. La pertinence de ce constat nous a été confirmée par un ancien évaluateur au ministère de la coopération : « *la participation active de la France au CAD débuta à la fin des années 1990 avec Jean-Louis Sabatier au ministère des affaires étrangères. Avant, le CAD et ses normes avaient peu d'influence sur l'évaluation de l'APD en France. Il aura donc fallu dix ans à la France pour s'approprier les cinq standards du CAD* »[358]. En France, le premier guide méthodologique sur l'évaluation de l'APD, publié en 1997 par le ministère des affaires étrangères, reflétait largement la pensée du CAD[359].

Cette tendance à la décentralisation dans la gestion de l'aide au développement se retrouve aussi au Royaume-Uni. En 1997, le CAD-OCDE félicita ce mode de fonctionnement du DFID[360] dans une évaluation par les pairs : « le Royaume-Uni a mis en place un système de gestion décentralisée de ses programmes d'aide au développement. C'est un outil efficace pour travailler en partenariat avec les gouvernements bénéficiaires, leur société civile et les autres

[355] *Ibid*.
[356] DAC-OECD, 2000, *op. cit.*
[357] Pour une analyse plus approfondie de ce phénomène, voir : TOULEMONDE, Jacques, "Europe and the Member States: Cooperating and Competing of Evaluation Grounds", in RIEPER, Olaf, TOULEMONDE, Jacques, *Politics and Practices of Intergovernmental Evaluation*, NJ: Transaction Publishers, 1997, pp. 117-132 ; SUMMA, Hikka, "Evaluation in the European Union: Addressing Complexity and Ambiguity", in FURUBO, Jean-Eric., RIST, ray, SANDHAL, Rolf, *International Atlas of Evaluation*, NJ: Transaction Publishers, 2002, pp. 407-424.
[358] Il fait référence aux premiers critères élaborés en 1986.
[359] Entretien avec un ancien évaluateur au ministère de la coopération français, consultant auprès d'organisations internationales, 4 avril 2013.
[360] L'ODA est devenue le DFID en 1997, lorsque le département a été séparé du Foreign and Commonwealth Office.

donateurs »[361]. La décentralisation apparaissait donc, à cette époque, comme une réforme préliminaire à des pratiques de gestion plus participatives.

Les agences d'aide engagèrent également des réformes pour passer d'une gestion par les moyens à une gestion par les résultats et partagèrent leurs expériences de cette réforme au sein du CAD-OCDE, afin que les donateurs les plus avancés en la matière guident leurs pairs[362]. Cela montre le caractère généralisé de la montée en puissance de la « bonne pratique » de la gestion axée sur les résultats dans les années 1990. Les États anglo-américains, plus imprégnés des principes du New Public Management, furent les premiers à instaurer des systèmes de gestion par les résultats. Tout d'abord, l'USAID élabora en 1994, *le Result-Based Management Framework*. En 1996, l'ACDI au Canada créa *une Result Based Management Policy* et le DFID instaura en 1998 un système de *Performance-Based Management*. Concernant la France, le CAD-OCDE loua la proximité entre le mode de gestion de l'AFD et celui des organismes du secteur privé : « l'autre force de l'AFD est son personnel, qui est recruté et géré selon les règles du secteur privé. Le personnel combine des profils professionnels divers, qui sont souvent considérés comme étant compétents et expérimentés »[363]. Ce commentaire du CAD est, de manière flagrante, imprégné de la pensée du New Public Management cherchant à rapprocher la sphère publique du secteur privé. Il reflète également les incitations de l'organisme pour que les donateurs bilatéraux adoptent des principes de gestion centrés sur l'évaluation de leur performance et de leur efficacité organisationnelle. Toutefois, l'approche de l'évaluation diffusée par le New Public Management tend à la faire se confondre avec la mesure des résultats, c'est-à-dire à la limiter à un objectif de redevabilité et d'évincer les effets d'apprentissage qu'elle peut apporter.

En France, la fonction d'évaluation était (et est toujours) divisée entre les trois organismes chargés de l'aide au développement. Le budget consacré à l'évaluation par ces organismes était de 1,5 millions de dollars. Au ministère des affaires étrangères, l'unité d'évaluation comptait sept membres, ils étaient cinq à l'AFD et deux au Trésor[364].

Au Trésor, la diffusion de l'évaluation fut facilitée par l'introduction de procédures de suivi dans la gestion des projets d'aide dans les années 1990. À cette époque, le Trésor commença également à sortir d'une logique d'évaluation de projets et se mit à réaliser des

[361] DAC-OECD, *Development Cooperation Review: United Kingdom*, Paris, 1997.
[362] DAC-OECD Working Group on Aid Evaluation, *Result based management in development aid agencies: a review of experiences*, 1999.
[363] DAC-OECD, 2000, *op. cit.*
[364] *Ibid.*

évaluations sectorielles, dans la santé et l'éducation principalement, ainsi que des évaluations de « stratégies pays ». La fonction d'évaluation s'élargit donc à un spectre plus large d'activités. L'organisme mit en place un manuel de standards d'évaluation pour se conformer à la bonne pratique prônée par le CAD-OCDE.

Au ministère de la coopération, l'évaluation s'institutionnalisa au fil de la décennie après un recul de la pratique à la fin des années 1980. Un évaluateur témoigne : « *la poussée de l'évaluation au ministère de la coopération prit fin en 1986 à cause de l'alternance politique avec la droite. Les périodes gouvernementales « à droite » ont toujours été défavorables à l'essor de l'évaluation. Il y eut une reprise en 1991 avec Patrice Tranchant (qui fut responsable des évaluations au ministère pendant plusieurs années). Il mit en place une unité d'évaluation un peu flottante. Elle était composée de trois personnes et était rattachée au cabinet du ministre de la coopération. Puis cette unité d'évaluation s'est institutionnalisée progressivement jusqu'en 1998, date du rattachement au ministère des affaires étrangères* »[365]. Cette unité fut rebaptisée Mission d'étude, d'évaluation et de prospective (MEPP) en 1998. Elle était chargée de relancer la culture de l'évaluation au sein du ministère de la coopération. Nous voyons dans cet historique que la fonction d'évaluation chercha sa place dans le ministère durant toute la décennie. En 1996, le ministère décida, lui aussi, de créer son propre guide méthodologique sur l'évaluation. Ce guide avait vocation à cadrer le mandat et les méthodologies de la MEEP et d'aligner les pratiques du ministère sur celles des organisations internationales (OCDE et Banque mondiale)[366]. Les projets étaient évalués en interne par les unités géographiques concernées et non par la MEPP, ce qui répondait à une logique de décentralisation et de responsabilisation du personnel. La MEPP se consacrait aux évaluations de plus large échelle : les évaluations sectorielles, les évaluations sur un pays, les évaluations d'outils et les méta évaluations.

Nous constatons que la fonction d'évaluation était organisée de manière assez semblable au Trésor et au ministère de la coopération. Ainsi, elle se développa et se structura sous l'influence forte des principes du New Public Management, très présents en Europe et dans les organisations multilatérales.

L'AFD avait la fonction d'évaluation la plus structurée. L'unité d'évaluation était située dans le département de l'audit général. Ce positionnement eut une influence importante sur la

[365] Entretien avec un ancien évaluateur du ministère de la coopération français, consultant auprès d'organisations internationales, entretien téléphonique, 4 avril 2013.
[366] *Ibid.*

perception de l'évaluation. Il donna le signal que la fonction de redevabilité de l'évaluation prenait le pas sur celle d'apprentissage. Ce rapprochement avec la fonction de l'audit renforça également l'aspect de « contrôle » que pouvait avoir l'évaluation dans l'esprit du personnel de l'agence, ce qui a pu nuire à sa fonction de dialogue et d'apprentissage. Cette décision fut d'ailleurs considérée *a posteriori* comme une erreur[367]. Selon un ancien directeur du département d'évaluation de l'AFD, « *à partir de 1993, l'accent était délibérément placé du côté de l'indépendance et du contrôle des financements par l'évaluation. En contrepartie, la fonction d'apprentissage et le lien avec les stratégies étaient moins privilégiés. Par ailleurs, la redevabilité était essentiellement assurée en interne et auprès du Conseil de surveillance* »[368]. Ces propos illustrent bien le virage effectué par les agences d'aide vers une pratique gestionnaire de l'évaluation qui prend une fonction de contrôle des pratiques du personnel et de mesure des résultats.

L'unité d'évaluation rendait des comptes à la Direction générale de l'agence, comme les autres unités, ce qui posait un problème d'autonomie selon le CAD-OCDE. Toutefois, les évaluations étaient menées en toute indépendance par rapport aux unités opérationnelles. Le CAD-OCDE recommanda ainsi en 2000 que « l'unité d'évaluation rapporte directement aux tutelles financières de l'Agence »[369]. À l'AFD, les évaluations étaient réalisées en interne et les méthodes d'évaluation étaient consignées dans un guide de bonnes pratiques réalisé en 1994. Selon nous, le fait que chaque organisme chargé de l'APD produisit de son côté un guide d'évaluation témoignait d'un manque de partage, voire une concurrence, entre les unités d'évaluation.

Dans les années 1990, sous la direction d'Antoine Pouillieute[370], l'AFD fit de nombreux efforts pour se conformer aux standards internationaux de bonne gestion de l'aide, notamment en matière d'évaluation. Ce directeur donna comme objectif à l'unité d'évaluation d'évaluer systématiquement les projets deux à trois ans après leur achèvement. Il décida également de la mise en place d'un mécanisme de suivi de tous les projets, les rapports d'achèvement de projets, réalisés dans les six mois suivant leur achèvement. Ces rapports devaient faciliter le travail des évaluateurs en assurant une documentation continue des projets durant leur mise en œuvre. Ils servaient de base à la réalisation d'évaluations sectorielles, par exemple sur les transports

[367] Propos issus de discussions informelles avec des évaluateurs de l'AFD lors d'une mission d'observation participante en 2012 et 2013.
[368] Entretien avec un ancien directeur du département de l'évaluation de l'AFD, 21 décembre 2011.
[369] DAC-OECD, 2000, *op. cit.* p. 64.
[370] Antoine Pouillieute a été fonctionnaire du Trésor public avant d'intégrer l'ENA par le concours interne. Il est ensuite entré au Conseil d'État, a été le directeur adjoint du cabinet d'Édouard Balladur. En 1995, il a été nommé directeur général de l'AFD, poste qu'il occupa jusqu'en 2001.

urbains, l'eau et l'assainissement, ou encore l'énergie. Enfin, Antoine Pouillieute dota l'AFD d'un comité des évaluations, où toutes les divisions opérationnelles étaient représentées pour discuter des résultats et des suites à donner aux recommandations des évaluations. Le comité soumettait ses conclusions au conseil de supervision de l'AFD. Ce mécanisme était censé améliorer l'utilisation des évaluations aux niveaux opérationnel et politique. L'utilité des évaluations devint une préoccupation centrale des unités d'évaluation dans les années 2000. Antoine Pouillieute fut donc un acteur important dans l'institutionnalisation de la fonction de l'évaluation au sein de l'AFD. Il réalisa cette institutionnalisation sous la pression du CAD-OCDE et des tutelles financières de l'AFD. Il a choisi de l'axer sur la redevabilité et la mesure de la performance. Il semblait voir l'évaluation comme une obligation et comme un instrument de gestion plutôt que comme une opportunité d'accéder à une meilleure connaissance sur l'aide et à rendre plus égalitaires les relations avec les partenaires du Sud de l'AFD.

Suivant cette logique d'amélioration de la redevabilité, l'AFD réalisait des rapports publics annuels synthétisant les travaux d'évaluation de projets. Le rapport annuel de 1998 montrait que 54 % des projets étaient évalués comme satisfaisants ou très satisfaisants. Près de la moitié des projets mis en place étaient donc jugés insatisfaisants. Ce score montre que les évaluations avaient un regard critique sur les activités de l'agence, bien qu'encore réalisées en interne. Si les objectifs fixés par les projets étaient souvent atteints, leur durabilité faisait défaut, avec seulement 39 % d'entre eux jugés durables. Quant aux recommandations formulées dans les évaluations, environ la moitié était mise en pratique[371].

Bien que le CAD-OCDE recommande la transparence au sujet des évaluations, les organismes français d'aide conservaient un principe de confidentialité. Au Trésor, les rapports d'évaluation étaient réservés à l'usage interne. À l'AFD et au ministère des affaires étrangères, la diffusion externe était limitée. Cela reflète la croyance que l'évaluation n'était destinée qu'aux donateurs. En 1998, le ministère des affaires étrangères réalisa un bilan de ses évaluations qui montra que les recommandations étaient utilisées au niveau opérationnel, mais très peu au niveau décisionnel. Un évaluateur estime que « *si les opérationnels utilisaient les évaluations, c'est grâce à la mise en place de comités de pilotage*[372] *des processus d'évaluation qui intégraient les services opérationnels. Par ailleurs, la mauvaise prise en compte des évaluations par les décideurs était due à la faible légitimité des évaluateurs dans le système de l'aide* »[373]. En effet, le statut des évaluateurs dans les agences d'aide en France était encore

[371] *AFD, Rapport annuel d'évaluation*, 1998, document interne à l'AFD.
[372] Pour une définition de ce terme, voir lexique.
[373] Entretien avec un ancien évaluateur du ministère de la coopération, 4 avril 2013.

flou, leur profession n'étant pas réellement reconnue. De plus, nous observons que les évaluations traitaient peu de sujets stratégiques, comme par exemple l'effet des programmes d'aide sur la baisse de la pauvreté, limitant l'intérêt que pouvaient y porter les décideurs politiques.

Les évaluateurs français se sentaient mal à l'aise avec la tournure gestionnaire que prit la fonction d'évaluation dans les années 1990. Ils se positionnèrent alors comme les défenseurs des approches participatives de l'évaluation dans les instances multilatérales. Un évaluateur du ministère des affaires étrangères déclara par exemple à l'OCDE en 1998 : « l'expérience montre que la participation des acteurs locaux aux processus d'évaluation est bénéfique. Nous avons besoin de la participation des experts nationaux pour valider et critiquer les résultats et les recommandations des évaluations »[374]. Toutefois, la mise en œuvre d'évaluations de ce type était encore peu courante, au dire même de ces évaluateurs : « il est difficile d'impliquer les autorités nationales, elles changent souvent et ne connaissent pas les projets »[375].

Concernant le Royaume-Uni, l'aide au développement était à la charge d'un seul organisme qui centralisait donc la fonction d'évaluation. Le département de l'évaluation du DIFD comptait dix agents plus du personnel de soutien. L'équipe d'évaluation se composait principalement de consultants externes et la fonction d'évaluation était indépendante des départements opérationnels, reportant directement au directeur général des ressources. Le budget consacré à l'évaluation s'élevait à environ deux millions de livres par an dans les années 1990 contre 500 000 livres au début des années 1980[376]. Cette hausse budgétaire témoignait de l'importance grandissante donnée à l'évaluation.

À la fin des années 1990, les évaluateurs britanniques connurent une transformation importante de leur fonction. En effet, aussi bien les thématiques que le public des évaluations, changèrent à partir de 1997, date à laquelle le gouvernement de Tony Blair engagea une grande réforme de la politique d'aide au développement. Dès lors, les évaluations eurent comme mission de prouver la contribution du Royaume-Uni à la lutte contre la pauvreté et elles furent rendues publiques. Un ancien évaluateur du DFID fait part de ces évolutions : « *en septembre 1997, le nouveau gouvernement commandita au DFID une évaluation de la politique d'aide bilatérale du Royaume-Uni, afin de préparer un livre blanc intitulé Eliminating World Poverty, qui recentrait l'aide au développement sur l'élimination de la pauvreté. Les termes de référence*

[374] Réponse à un questionnaire du CAD-OCDE sur la pratique de l'évaluation dans les agences d'aide lancé en 1998.
[375] Réponse à un questionnaire du CAD-OCDE sur la pratique de l'évaluation dans les agences d'aide lancé en 1998.
[376] DAC-OECD, 1998, *op. cit.*

de cette évaluation étaient novateurs. Cette évaluation avait un objectif prospectif. Elle devait contribuer au débat politique et se dérouler sur deux années. C'était un défi important pour nous, évaluateurs, car le public de l'évaluation avait changé rapidement. Par ailleurs, c'était la première évaluation du DFID qui posait la question de savoir si les programmes d'aide faisaient une différence dans la vie des populations pauvres »[377]. Ce témoignage illustre bien le tournant que prit l'évaluation à la fin des années 1990 en analysant des enjeux plus globaux, comme la réduction de la pauvreté, et en visant à informer un public large, des décideurs politiques à la société civile.

Les dirigeants du DFID, conscients de l'importance croissante de la prise en compte des aspects sociaux du développement dans les évaluations des projets de lutte contre la pauvreté, cherchèrent à augmenter les capacités du DFID en termes d'analyse sociologique, notamment avec l'initiative *Social Development Systems for Coordinated Poverty Eradication* (SCOPE). Elle avait pour objectif d'augmenter la qualité des théories de développement social utilisées par le DFID, ainsi que les méthodes et les indicateurs pour améliorer les interventions de développement social.

La décision de rendre publics les rapports d'évaluation fit suite, selon nous, à l'adoption de la *Citizens' Chart* au début des années 1990, qui renforçait le principe de transparence des politiques publiques. Ainsi, les évaluations et leurs synthèses étaient généralement rendues accessibles au public, même lorsqu'elles critiquaient sur les actions évaluées. Cette transparence fut louée par le CAD-OCDE qui considérait qu'elle renforçait la crédibilité du système d'évaluation du DFID. D'après les bases de données du CAD[378], 23 évaluations ont été produites et publiées entre 1995 et 1997, ainsi que quatre études sectorielles. Nous voyons dans cette hausse de la redevabilité et de la transparence des évaluations du DFID, un reflet du modèle du New Public Management, selon lequel l'évaluation devrait, avant tout, servir à rendre compte à l'opinion publique des performances de l'administration.

Concernant la diffusion interne des évaluations, les évaluateurs du DFID font part d'évolutions importantes durant les années 1990. Un ancien évaluateur au DFID raconte qu'« *auparavant la diffusion des conclusions des évaluations ne nous souciaient qu'une fois l'évaluation achevée. Aujourd'hui, pour les évaluations stratégiques, nous essayons d'impliquer le personnel du DFID tout au long du processus évaluatif. La priorité des évaluateurs est de mettre en place une base pour la communication en amont de l'évaluation,*

[377] Entretien avec un ancien évaluateur du DFID, 17 juin 2013.
[378] DAC-OECD, 1997, *op. cit.*

pas seulement à la fin. Toutefois, cela a été ardu pour les évaluateurs. Les dirigeants étaient impatients et voulaient les résultats alors que l'évaluation avait à peine commencé »[379].

Selon les évaluateurs du DFID, le principe du CAD-OCDE qui fut le plus difficile à atteindre était celui de la pertinence des évaluations. Ils expliquent que « *les évaluations ex post stratégiques (sur un secteur d'activité ou sur un pays) sont déconnectées du temps politique. Les évaluations devraient être davantage prospectives et regarder vers l'avenir et non vers le passé pour fournir des leçons qui sont utiles aux enjeux émergents de l'aide* »[380]. De manière générale, les évaluateurs britanniques portaient un regard positif sur les critères du CAD-OCDE qu'ils estimaient correspondre à leur culture organisationnelle[381].

Ces constats comparatifs sur la France et le Royaume-Uni tendent à valider notre hypothèse de départ, selon laquelle la diffusion des standards internationaux d'évaluation ne peut s'effectuer que si la culture des politiques publiques au niveau national s'y prête. L'évaluation, bien qu'apparemment technique, est porteuse de valeurs et de conceptions du politique et elle ne peut être importée dans un système politique de manière neutre sans entraîner résistances et réappropriations. La culture nationale des politiques publiques a un poids significatif sur le degré et la manière dont les bonnes pratiques internationales, telles que l'évaluation, sont appropriées[382]. Cette analyse montre que le processus de normalisation de la fonction d'évaluation en était encore à ses débuts dans les années 1990. Les organismes d'aide issus de pays où la culture de l'évaluation était peu ancrée avaient des difficultés et/ou des réticences à se conformer aux principes du CAD-OCDE. En effet, plus que l'influence normative de cette institution, ce qui permet la diffusion d'une culture de l'évaluation dans les agences d'aide, c'est l'appropriation des principes de cette culture par les dirigeants politiques nationaux qui pourront en relayer l'esprit aux agents des administrations[383].

Même s'il est imparfait, le processus de refonte et de normalisation de la gestion de l'aide au développement selon les principes du New Public Management avait déjà des effets observables sur les activités d'aide. En effet, sous couvert de réformes de décentralisation et de hausse de l'autonomie des opérateurs, une bureaucratisation des agences d'aide eut lieu, avec

[379] Entretien avec un ancien évaluateur du DFID, 17 juin 2013.
[380] Entretien avec un ancien évaluateur interne de l'ODA, 14 janvier 2013.
[381] Entretien avec un ancien évaluateur au DFID, 17 juin 2013.
[382] KLEIN, Asmara, LAPORTE, Camille, SAIGET, Marie, *Les bonnes pratiques des organisations internationales,* Presses de Sciences po, à paraître en mars 2015.
[383] Dans la deuxième partie de la recherche, nous verrons plus en détail les jeux d'acteurs qui permettent la diffusion d'une culture de l'évaluation.

un afflux de procédures et de règles à suivre afin d'assurer un suivi et une évaluation continus des activités.

Reprenant les termes de Béatrice Hibou, nous estimons que les programmes d'aide au développement, qui étaient qualifiés d'« éléphants blancs »[384] dans les années 1970-1980, devinrent des « éléphants de papiers » dans les années 1990[385]. Les programmes étaient mis en place par des experts qui devaient suivre des procédures administratives très sophistiquées pour rendre compte à leur hiérarchie. Les opérateurs de projets eurent le sentiment de passer d'un statut « d'hommes de terrain » à un statut de « bureaucrates » devant la multiplication des rapports (de suivi, d'achèvement, d'évaluation, etc.) et la nécessité de continuellement collecter des données quantitatives[386]. Ils se retrouvèrent confrontés, tout comme leurs directions, à un processus de surenchère bureaucratique de dispositifs de contrôle, de mesure, d'indicateurs, d'évaluation, de transparence, etc. Cela amène à penser que l'augmentation des procédures de reddition de comptes, au lieu d'améliorer la transparence de l'aide, se traduisit par une hausse de son opacité (perçue).

Les « bonnes pratiques » de suivi, d'évaluation, de transparence et de redevabilité étaient censées donner une image rationnelle de l'aide au développement, tournée vers la réduction de la pauvreté. Si au niveau multilatéral, dans les sommets sur le développement, cette image rationnelle fonctionnait et était transmise à « l'opinion publique mondiale », nous considérons que les bailleurs de fonds ne mettaient en place que les standards qui correspondaient le mieux à leur culture organisationnelle, délaissant les autres. La multiplication des bonnes pratiques de gestion de l'aide rendait plus difficile leur contrôle et donnait un argument aux bailleurs de fonds pour ne pas les appliquer. Pour appuyer cet argument, nous pouvons nous référer aux termes de Béatrice Hibou : « le processus normatif est si tâtonnant et approximatif, les bailleurs changent si souvent de stratégie et de mode, les techniciens sont si friands d'améliorations, les décideurs sont si demandeurs d'opérationnalité et de meilleure appréhension du réel, les évaluations sont si fréquentes, les innovations sont si nombreuses que l'on assiste à une véritable jungle bureaucratique »[387].

[384] L'expression fait référence à des programmes de développement de grande envergure qui sont noyautés par des exigences bureaucratiques qui rendent leurs mises en œuvre très lourde pour les partenaires locaux et les responsables de programmes des organismes d'aide.
[385] HIBOU, 2012, *op. cit.* p. 73.
[386] Constat fait à partir de conversations informelles avec des agents « séniors » de l'AFD lors de notre observation participante de 2012 à 2013.
[387] HIBOU, 2012, *op. cit.* p. 91.

- **Conclusion de chapitre**

Les années 1990 marquèrent donc le début de la diffusion de bonnes pratiques internationales de l'évaluation dans le système de l'aide au développement. Ces bonnes pratiques furent déterminées en cercle fermé par les donateurs occidentaux au sein du CAD-OCDE sous l'influence politique des donateurs anglo-américains et du Nord de l'Europe, dont les politiques publiques connaissaient une vague de réformes inspirée du New Public Management.

Durant les années 1990, l'évaluation oscilla entre, d'une part une fonction d'apprentissage sur le développement, avec des analyses menées sur les nouveaux enjeux du développement (genre, développement durable, bonne gouvernance) ou selon des méthodes participatives et, d'autre part une fonction gestionnaire (mesure des résultats, redevabilité). Dans un contexte où les organismes d'aide avaient besoin de relégitimer leur mandat, c'est la première fonction qui l'emporta. Pour paraître redevables et performantes les organismes d'aide se devaient d'évaluer systématiquement leurs projets et programmes d'aide selon des standards et des procédures harmonisés. Loin d'être purement techniques, ces bonnes pratiques, étaient porteuses d'une approche politique particulière de l'aide qui donnait une place centrale à l'obtention de résultats à court terme, à la transparence et à la prise de décision rationnelle fondée sur des preuves.

L'institutionnalisation de l'évaluation dans les organismes d'aide, selon des standards spécifiques, entraîna en retour des changements d'ordre politique au sein des organismes de développement, point qui sera exploré en filigrane tout au long de notre recherche.

Chapitre 3 : 2000-2011, l'évaluation, un outil de dépolitisation de l'aide au développement

Le passage au 21ème siècle a été l'occasion pour les bailleurs de fonds de se lancer dans une ère de renouveau. Nous avons choisi de débuter ce chapitre par l'année 2000, car elle a été marquée par l'adoption des Objectifs du Millénaire pour le Développement (OMD) qui sont devenus le référentiel universel à partir duquel les acteurs de l'aide au développement évaluent leurs performances. Ensuite, en 2005, la Déclaration de Paris a été adoptée, instituant un ensemble de bonnes pratiques auxquelles les donateurs et les gouvernements bénéficiaires doivent se conformer pour gérer efficacement l'aide au développement et atteindre les OMD. Cette déclaration a, dès lors, constitué le référentiel d'évaluation de la mise en œuvre de l'aide au développement. Ainsi, la pratique de l'évaluation a été normalisée dans ses procédures (avec les principes du CAD-OCDE de 1991) et dans son contenu (« bonnes pratiques » de la Déclaration de Paris et OMD).

Dans ce chapitre, nous allons démontrer que l'évaluation est devenue, dans les années 2000, un outil de contrôle permettant de garantir que les politiques des donateurs contribuent bien aux OMD et respectent les « bonnes pratiques » de la Déclaration de Paris. Nous notons également un retour aux évaluations expérimentales populaires dans les années 1960, que nous identifions comme une volonté de dépolitiser l'évaluation. En tant qu'instrument scientifique, l'évaluation permet ainsi de légitimer les réformes apportées par les OMD et de la Déclaration de Paris, en prouvant objectivement leurs effets positifs sur les résultats de l'aide au développement. Enfin, si la pratique de l'évaluation apparaît harmonisée entre les donateurs, nous montrerons qu'une diversité d'approches et de pratiques persistent, reflétant les cultures politiques des États donateurs. Ce chapitre historique s'arrête en 2011, date du Haut Forum de Busan qui marque la fin du « régime de l'efficacité de l'aide » qu'a incarné la Déclaration de Paris[388].

[388] Certains des enjeux qu'a entraînés ce Forum en matière d'évaluation seront abordés dans le chapitre 7 de la thèse.

A. Des approches contradictoires de l'évaluation

Nous allons, tout d'abord, démontrer que les référentiels des OMD et de la Déclaration de Paris présentent des approches de l'évaluation qui peuvent être contradictoires, entre lesquelles les donateurs et les évaluateurs doivent arbitrer dans la pratique. Cette situation entraîne un certain flou sur les finalités de l'évaluation et l'adoption d'approches « hybrides » de l'évaluation par les donateurs qui « piochent » dans les deux référentiels selon leurs orientations politiques et leurs besoins organisationnels[389].

- **Les Objectifs du Millénaire : une approche « équitable » de l'évaluation**

Les OMD[390] symbolisent une obligation de performance pour les acteurs de l'aide au développement et font l'objet d'un suivi régulier. Les acteurs de l'aide doivent rendre compte régulièrement aux Nations unies de leurs progrès vers l'atteinte de ces objectifs. Les OMD ont été adoptés pour donner une finalité universelle à l'aide au développement et pour améliorer l'éthique et l'efficacité de cette dernière. Selon les OMD, l'aide doit être allouée aux populations les plus pauvres[391] et avoir des résultats rapides et concrets. Le cadre des OMD oscille donc entre une approche de l'aide comme un droit universel envers les populations les plus pauvres, quelles que soient les capacités de leurs gouvernements, et une approche de l'aide conditionnée par l'obtention de résultats chiffrés, rapides et continus. Nous allons démontrer que malgré le discours universaliste et égalitaire qui entoure les OMD, l'aide au développement répond toujours, dans les années 2000, à une logique de conditionnalité et de réduction des risques. L'évaluation reste un outil de redevabilité plus qu'un outil d'apprentissage et les tendances observées dans les années 1990, en termes de gestion néolibérale de l'aide, s'amplifient.

La réflexion des donateurs sur la nécessité d'établir un ensemble d'objectifs mesurables pour relégitimer l'aide au développement a débuté, à la fin des années 1990 au sein du CAD-OCDE, à la suite de la publication du rapport *Shaping the 21st century : the role of development*

[389] SASAKY, Ryoh, "A Review of the History and the Current Practices of Aid Evaluation", *Journal of Multidisciplinary Evaluation,* n°5, 2006, pp. 55-83.

[390] Les OMD sont au nombre de huit : réduire l'extrême pauvreté et la faim, assurer à tous l'éducation primaire, promouvoir l'égalité des genres et l'autonomisation des femmes, réduire la mortalité infantile, améliorer la santé maternelle, combattre le VIH/SIDA, le paludisme et les autres maladies, assurer un environnement humain durable et construire un partenariat mondial pour le développement. Ils sont divisés en 18 indicateurs quantitatifs de suivi qui permettent de mesurer les progrès des acteurs de l'aide vers leur atteinte d'ici à 2015.

[391] Avec ici une conception multidimensionnelle de la pauvreté qui dépasse la condition économique des individus pour englober leurs conditions de vie sanitaire, éducative et sociale.

cooperation[392]. Les représentants du Japon ont introduit l'idée qu'« un nombre limité d'objectifs quantitatifs devrait être établi pour rendre compte des résultats du développement sur le moyen terme, afin que la coopération pour le développement soit perçue comme contribuant à un changement perceptible dans les pays en développement »[393]. Le CAD a décidé que les objectifs seraient déterminés de façon sectorielle, reprenant les grandes thématiques développées par les Nations unies durant les années 1990 (lutte contre la pauvreté, environnement, genre, etc.). Les représentants de l'Allemagne ont fait un plaidoyer vigoureux en faveur de l'inclusion d'un objectif portant sur les enjeux de gouvernance et des Droits de l'Homme, mais leur proposition n'a pas été retenue, notamment en raison d'une opposition américaine et japonaise[394]. Les négociations finales sur ces objectifs ont eu lieu au sein des Nations unies à l'occasion de la préparation de l'Assemblée générale de l'année 2000. Les OMD sont donc le fruit d'un compromis entre les grandes organisations multilatérales : les Nations unies, l'OCDE, ainsi que la Banque mondiale et le FMI qui ont été associés aux négociations[395]. Ils ne constituent donc pas un consensus de la « communauté du développement » dans son ensemble, l'accord n'incluant ni les bénéficiaires de l'aide, ni les organisations de la société civile, ni les acteurs privés du développement et ni les donateurs non-membres du CAD-OCDE.

La phase finale de négociation a entraîné la suppression des objectifs portant sur la santé reproductive et sexuelle, excepté en ce qui concerne le sida. Les OMD n'ont donc pas mis l'accent sur de nouvelles priorités de développement par rapport aux années 1990. Le renouveau se situait non pas dans les thématiques abordées, mais dans le fait de donner un ensemble d'objectifs universels à l'aide au développement, à partir desquels son efficacité pourrait être évaluée de manière transparente. Selon Richard Manning, ancien président du CAD-OCDE, « les OMD ont renforcé le poids de la mesure des résultats dans l'aide au développement. Ils ont fortement encouragé les pays bénéficiaires de l'aide et la communauté des donateurs à être plus sérieux quant à ce que leurs programmes accomplissaient réellement »[396].

Après une décennie moribonde, donner des objectifs universels, simples et mesurables à l'aide au développement revenait à en véhiculer une image rationnelle, cohérente et

---

[392] DAC-OECD, 1996, *op. cit.*
[393] *Ibid.*
[394] Archives du CAD-OCDE, compte rendu des réunions de 1996 à 1997 sur les suites du rapport *Shaping the 21st Century* de 1996.
[395] HULME, David, *the Making of the Millennium Development Goals: Human Development Meets Result-Based Management in an Imperfect World*, University of Manchester, Institute for Development Policy and Management, 2007.
[396] MANNING, Richard, *The DAC as a central actor in development policy issues: experiences over the past four years*, Discussion Paper, Deutsches Institut für Entwicklungspolitik, Bonn, 2008.

transparente. Ces objectifs et indicateurs sont également apparus comme des instruments permettant une évaluation globale de l'efficacité de l'aide, problème qui mobilisait et divisait les experts du développement depuis le milieu des années 1980. Selon Jean-Michel Severino, ancien directeur général de l'AFD, « les OMD sont des objectifs simples, concrets, explicables, autour desquels on peut communiquer et qui permettent surtout de confronter les résultats aux anticipations. Ils représentent pour beaucoup d'organisations une chance de démontrer leur contribution aux objectifs globaux de réduction de la pauvreté (…) Non seulement ces objectifs sont utiles en termes quantitatifs, mais ils nourrissent un dialogue sur la nature des politiques publiques à mener pour les atteindre, sur la pertinence des objectifs eux-mêmes, sur les relations entre les moyens financiers et les résultats, etc. »[397].

Les OMD remplissent des fonctions multiples : la mesure de la performance, la légitimation de l'aide, l'accumulation de connaissances (en constituant une base de données internationale sur le développement des pays du Sud) et la communication sur l'aide au développement. Toutes ces qualités ont fait qu'ils sont rapidement devenus le standard international d'évaluation de l'efficacité de l'aide. Au début des années 2000, les évaluateurs se sont donc retrouvés avec plusieurs référentiels normatifs d'évaluation à respecter : les critères du CAD-OCDE sur la qualité des évaluations et les OMD sur le contenu des évaluations. Nous soulignons le fait que les OMD ne servent pas seulement à mesurer la performance des donateurs, mais également des gouvernements des États du Sud. Ils tendent ainsi implicitement à « responsabiliser » les bénéficiaires de l'aide, dans l'esprit néolibéral de la décennie précédente[398].

Les OMD ont une portée normative forte, dans la mesure où ils guident le contenu des programmes d'aide et des évaluations. Les thématiques qui ne sont pas incluses dans les OMD (Droits de l'Homme, santé reproductive, renforcement des institutions, etc.) sont laissées de côté par les organismes d'aide. Ces objectifs placent également l'efficacité globale de l'aide en haut de l'agenda politique des organisations multilatérales et bilatérales. Ces dernières doivent impérativement produire des résultats continus.

Les OMD sont présentés comme un outil de suivi de l'efficacité de l'aide simple et quantifiable. Selon nous, ce sont effectivement de bons outils pour communiquer clairement sur l'aide au développement, mais ils ne simplifient aucunement l'évaluation de son efficacité. Premièrement, ils agrègent des indicateurs multiples, dont la pondération est délicate sachant

[397] SEVERINO, Jean-Michel, « Vers des objectifs mondiaux pour la globalisation », CERISCOPE *Pauvreté*, 2012.
[398] Les OMD ont été ratifiés par 189 États.

qu'ils sont souvent interdépendants. Deuxièmement, les résultats des évaluations varient selon qu'elles ont comme champ d'analyse un pays ou bien un groupe de population (un pays très inégalitaire peut avoir rempli un des OMD car une partie de la population se développe rapidement, alors qu'une autre partie stagne ou régresse), ou selon qu'elles présentent des résultats agrégés (de l'objectif dans son ensemble) ou désagrégés (indicateur par indicateur), un pays pouvant faire des progrès globaux en matière de santé infantile grâce à un taux élevé de vaccination, mais qui cache des structures sanitaires en déliquescence[399]. De plus, le suivi et l'évaluation des OMD nécessitent la constitution et le renseignement, sur plusieurs années, de bases de données statistiques sur des thèmes pour lesquels les données sont parcellaires dans les pays du Sud, notamment le genre.

Avec l'adoption des OMD, l'aide au développement est devenue sujette à un suivi-évaluation constant qui a lieu à tous les échelons du système de l'aide : au niveau multilatéral, un suivi annuel est réalisé par les Nations unies, qui donne lieu à des conférences très médiatisées ; au niveau bilatéral, les agences d'aide évaluent leur performance à l'aune des OMD ; au niveau national, la plupart des pays en développement se servent des OMD comme référentiel d'évaluation de leurs politiques publiques. Par exemple, les bureaux nationaux du PNUD ont assisté les pays bénéficiaires de l'aide des Nations unies dans l'élaboration de documents de planification du développement national qui intègrent les indicateurs des OMD[400].

Les OMD marquent le renouveau du principe d'universalité de l'aide des années 1960, bien qu'il soit désormais teinté de l'idée d'efficacité. Si aucun groupe de population n'atteignant pas les OMD ne doit être exclu de l'aide, cette dernière doit être efficace et les populations doivent se donner les moyens d'améliorer leur état de développement. Ce mélange entre universalisme et rationalisme devait susciter un élan des budgets d'aide dans les pays du Nord. La stratégie des Nations unies a fonctionné, dans la mesure où nous observons une hausse de 63 % de l'APD mondiale nette entre 2000 et 2010[401]. Nous observons donc une corrélation entre l'adoption des OMD et la hausse des volumes d'aide au développement à partir du début des années 2000.

L'approche des OMD est ambiguë : si les pays pauvres doivent se responsabiliser pour lutter contre la pauvreté, ils ne sont pas responsables de cet état. Selon Jean-David Naudet,

[399] SEVERINO Jean-Michel, 2012, *art. cit.*
[400] STOKKE, Olav, 2009, *op. cit.*
[401] Statistique du CAD-OCDE http://stats.oecd.org/qwids/, consulté le 2 septembre 2013.

« avec les OMD, un retour est effectué vers une analyse situationnelle de la pauvreté dont les causes sont en dehors de la responsabilité des gouvernements et des peuples du Sud »[402]. Les OMD rompent, en effet, pour un temps, avec le discours selon lequel la pauvreté serait le résultat de la mauvaise gouvernance des États. Pour appuyer ce tournant idéologique sur des « preuves scientifiques », les Nations unies ont commandité une étude sur la pauvreté à l'économiste américain Jeffrey Sachs. Il analyse que la pauvreté relève de la situation géographique des pays, mais surtout de la responsabilité des pays du Nord et de leurs politiques agricoles envers les pays en développement[403]. Par ailleurs, en raison du phénomène de « trappe à pauvreté »[404], la mauvaise gouvernance serait une conséquence et non une cause du sous-développement. L'aide retrouve ainsi sa légitimité en tant que réparation morale des donateurs. Jeffrey Sachs estime également que la réalisation des OMD nécessite une augmentation de l'aide au développement qui correspond, par une coïncidence confondante, à 0,7 % du PNB des donateurs, objectif perpétuellement fixé depuis les années 1960.

La politique de l'aide au développement dans les années 2000 regroupe les idées d'un impératif moral et d'un impératif d'efficacité. Aujourd'hui, une aide morale mais qui n'apporte pas de résultats n'est pas légitime. Ainsi, l'évaluation, en prouvant l'efficacité des pratiques et des politiques d'aide, est un instrument de légitimation de l'aide au développement. Alors qu'avant les années 1990, les donateurs avaient seulement besoin de se légitimer lorsque leurs pratiques faisaient l'objet de fortes critiques, après des décennies de remise en cause de l'efficacité de l'aide, ce besoin de légitimation est désormais constant.

Les deux impératifs, moral et d'efficacité, de l'aide au développement modifient les déterminants de la conditionnalité de l'aide. Étant donné que les OMD ont une visée universelle pour permettre une égalité des chances entre les individus quelles que soient leurs conditions d'origine, il est moins question de discriminer les montants d'aide en fonction de la gouvernance des pays[405]. D'autant plus que la mauvaise gouvernance est désormais considérée comme une conséquence de la pauvreté. La « bonne pratique » en matière de sélectivité est désormais d'allouer l'aide en fonction d'objectifs fixés *ex ante* aux pays à partir d'une évaluation de leurs besoins et de leurs capacités réalisée conjointement par les donateurs et les

[402] NAUDET, Jean-David, 2005, *art. cit.*
[403] SACHS, Jeffrey, *Investir dans le développement : plan pratique pour réaliser les Objectifs du Millénaire pour le développement*, Millenium Projet, 2005.
[404] En 1943, Paul Rosenstein-Rodan, a montré qu'un pays peut se trouver piégé dans une "trappe à pauvreté" lorsque les investissements industriels dont il bénéficie ne trouvent pas de relais dans des activités complémentaires. En l'absence d'effets d'entraînement, un pays peut ainsi être enfermé dans un cercle vicieux de sous-développement, quelles que soient ses tentatives d'industrialisation.
[405] Cette idée sera rapidement abandonnée comme nous le verrons par la suite.

gouvernements du Sud. Or, il est très difficile d'évaluer ces deux aspects. Selon Jean David Naudet, « souvent, ni les gouvernements partenaires, ni les donateurs n'ont la capacité d'évaluer les capacités et les besoins globaux au niveau d'un pays. De plus, il n'existe pas de critères standardisés d'évaluation de ces capacités et de ces besoins au niveau international, ce qui empêche une sélectivité harmonisée entre les bailleurs de fonds »[406]. Une autre difficulté pour les évaluateurs est de convaincre les gouvernements de les laisser accéder aux populations bénéficiaires pour que celles-ci fassent part de leurs besoins et de leurs difficultés[407].

Au-delà de cet aspect, les OMD ont élargi le champ de l'évaluation de l'aide au développement. Un des principes fondateurs de la Déclaration du Millénaire est que l'aide doit rétablir l'équité entre les groupes sociaux les plus démunis et les plus aisés. L'évaluation doit, en conséquence, vérifier que la mise en œuvre des activités d'aide a bien été équitable et que leurs résultats améliorent l'équité au sein des sociétés, notamment en matière de genre. Les évaluateurs, de l'UNICEF et d'ONU Femmes en particulier, ont ainsi développé une méthodologie pour mettre en place des protocoles d'évaluation mesurant l'équité des programmes d'aide : les *equity based evaluations*[408]. Les évaluateurs de ces organismes ont diffusé cette méthodologie auprès de l'ensemble des évaluateurs de l'aide au développement, en organisant des séminaires internationaux retransmis sur Internet et en diffusant des guides de bonnes pratiques[409]. La prise en compte de l'équité, notamment en matière de genre, est ainsi devenue un nouveau standard qualitatif de l'évaluation. Aujourd'hui, même les donateurs les moins impliqués sur le genre développent des procédures sur la prise en compte du genre dans les programmes d'aide et leurs évaluations. Le ministère français des affaires étrangères a, par exemple, développé une « stratégie genre et développement »[410] pour le programme quinquennal 2013-2017.

Les OMD présentent la pauvreté comme un enjeu multidimensionnel ne pouvant être traité que par des interventions simultanées portant sur l'ensemble des dimensions qui le compose (emploi, santé, éducation, gouvernance, environnement, égalité, etc.). Ils remettent donc en cause la pertinence de l'aide au développement sous la forme de projets indépendants et favorisent une approche en termes de programmes sectoriels ou multisectoriels, liant par exemple la santé et l'éducation. Le budget requis pour le financement de ces larges programmes

[406] NAUDET, Jean-David, 2005, *op. cit.*
[407] LAPORTE, Camille, *Évaluation et efficacité de l'aide au développement en Corée du Nord*, L'Harmattan, 2011, 220p.
[408] UNICEF, *How to design and manage equity-focused evaluations,* 2011; UNIFEM, *UNIFEM works on gender-responsive budgeting,* Evaluation Unit, 2009.
[409] Par exemple, http://mymande.org/equity_focused_evaluation, consulté le 15 décembre 2013.
[410] Ministère des affaires étrangère, *Stratégie genre et développement 2013-2017,* Direction générale de la mondialisation, 2013.

étant plus élevé que pour celui de projets de petite échelle, les bailleurs de fonds se regroupent souvent pour réunir le budget nécessaire. Il devient alors de plus en plus complexe pour les évaluateurs de répondre à la demande des tutelles financières des agences d'aide de mesurer leur contribution particulière au développement des pays bénéficiaires. La solution serait que les dirigeants des agences d'aide abandonnent cette vision individuelle de l'efficacité et qu'ils systématisent la pratique des « évaluations conjointes »[411]. Toutefois, les exigences de redevabilité qui pèsent sur les agences d'aide semblent empêcher cette évolution[412].

Les donateurs allouent également leur aide au développement sous la forme de soutien budgétaire[413] aux pays bénéficiaires, c'est-à-dire que les gouvernements sont libres d'utiliser les fonds selon leur propre stratégie de développement (à condition de rendre compte de cette utilisation aux donateurs). Dans ce cas, le travail des évaluateurs des organismes d'aide n'est plus vraiment d'évaluer les programmes de leur organisme, mais d'évaluer les politiques publiques des pays bénéficiaires. Selon un ancien directeur de la division de l'évaluation et de la capitalisation à l'AFD, « *il y a toujours une ambiguïté, un flou, dans la distinction entre l'évaluation du projet et l'évaluation de la politique publique du pays bénéficiaire. Il ne faut pas que l'évaluation fasse le jugement des politiques publiques des pays bénéficiaires, mais souvent c'est le cas de manière inhérente* »[414]. Or, le fait que des évaluateurs étrangers évaluent leurs politiques publiques peut être perçu comme de l'ingérence par les pays bénéficiaires. Une solution serait que les pays bénéficiaires évaluent eux-mêmes leurs politiques de développement et transmettent ces évaluations aux donateurs, mais comme nous le verrons dans la suite de cette recherche, les obstacles sont multiples à cette évolution.

Ces nouvelles formes de l'aide au développement transforment et complexifient donc la tâche des évaluateurs[415]. Il devient difficile pour eux d'attribuer les résultats de l'aide à des actions précises et à des donateurs individuels. Une autre complexification provient des contradictions entre le référentiel d'évaluation des OMD et celui de la Déclaration de Paris qui vient s'y ajouter en 2005.

[411] Les évaluations conjointes sont des évaluations réalisées par une équipe regroupant plusieurs donateurs sur un secteur, un pays ou un programme commun.
[412] Constat réalisé lors de notre observation participante à l'AFD et lors de séminaires sur l'évaluation regroupant des donateurs européens.
[413] Pour une définition de la notion « d'aide budgétaire » voir lexique en annexe 7.
[414] Propos recueillis lors de conversations informelles lors de notre observation participante à l'AFD de 2012 à 2013.
[415] CONLIN, Sean, STIRRAT, Roderick, "Current challenges in development evaluation", *Evaluation,* vol. 14, n°193, 2008, pp. 193-208.

- **La Déclaration de Paris : un usage gestionnaire de l'évaluation**

Si les OMD sont devenus, dans les années 2000, le référentiel international de l'évaluation des résultats de l'aide au développement, un autre référentiel s'est également imposé, celui-ci cherchant à instituer un ensemble de bonnes pratiques en matière de gestion de l'aide : la Déclaration de Paris. Ces deux aspects, les résultats et la gestion de l'aide, vont en effet de pair. Il est impossible d'atteindre de bons résultats, les OMD, si la gestion de l'aide n'est pas elle-même efficace. Pour marquer le passage à une nouvelle ère du développement, il ne suffisait donc pas d'établir un ensemble d'objectifs harmonisés internationalement, les donateurs devaient également montrer qu'ils changeaient les instruments de mise en œuvre de l'aide et qu'ils harmonisaient leurs pratiques.

Une série de conférences internationales ayant pour objet d'établir un ensemble de bonnes pratiques de gestion de l'aide au développement a eu lieu au début des années 2000. En 2002, l'Assemblée générale des Nations unies, présidée par Kofi Annan, a adopté le Consensus de Monterrey[416] portant sur les bonnes pratiques à adopter en matière de financement du développement[417].

Suite à cet accord international, le CAD-OCDE a organisé, à intervalle régulier, des Forums de Haut Niveau sur l'efficacité de l'aide, afin d'établir les principes internationaux d'une gestion plus efficace de l'aide. Selon le coordinateur du groupe de travail sur l'efficacité de l'aide du CAD, « l'engagement fort des bailleurs de fonds, et surtout des ministère des finances, dans les Forums de haut niveau sur l'efficacité de l'aide s'explique par leur volonté de se réapproprier et de mieux contrôler l'aide au développement dont les flux augmentent de manière importante »[418]. Nous retrouvons ici l'idée forte chez les donateurs de contrôler davantage l'aide au développement : finalité que l'évaluation va servir. En 2003, un Forum de Haut Niveau a été organisé à Rome regroupant 150 représentants de pays donateurs et bénéficiaires de l'aide. Il a abouti à des engagements des donateurs en matière d'harmonisation de leurs pratiques de gestion de l'aide et de suivi chiffré de leurs résultats[419]. En 2005, un deuxième forum a été organisé, donnant lieu à la Déclaration de Paris[420] qui énonce un ensemble

[416] Assemblée générale des Nations unies, 2002, *op. cit.*
[417] Les donateurs s'engagèrent à augmenter leurs fonds d'aide au développement pour atteindre la cible des 0,7 %, à harmoniser leurs pratiques, à rendre plus prévisibles leurs allocations d'aide, à prendre en compte les capacités d'absorption des pays bénéficiaires, à s'aligner sur les cadres de développement des pays bénéficiaires, à adopter une gestion axée sur les résultats, ainsi qu'à travailler plus en partenariat avec les acteurs du secteur privé.
[418] Hubert de Milly, allocution à la Société française d'évaluation, jeudi 16 octobre 2008, cahier de la SFE, *op cit.*
[419] DAC-OECD, *Harmonizing donor practices for effective aid delivery*, 2003.
[420] DAC-OECD, *Déclaration de Paris sur l'aide au développement,* 2005. Les objectifs de la Déclaration de Paris sont :

de cinq principes sur la gestion de l'aide[421], dont notamment la gestion axée sur les résultats[422]. Ce document est devenu le pilier normatif de la gestion de l'aide et les Forums de haut niveau ont, dès lors, été consacrés à son suivi et au perfectionnement de ces principes.

Toutefois, les négociations de la Déclaration de Paris ont été largement critiquées par les ONG internationales qui a condamné le fait qu'un « club de donateurs » fermé produise des bonnes pratiques engageant l'ensemble de la « communauté de l'aide »[423]. Pourtant, Richard Manning, président du CAD-OCDE de 2003 à 2007, avait à cœur de rendre le forum de Paris plus inclusif que celui de Rome. Richard Manning, ancien directeur exécutif à la Banque mondiale et manager au DFID, est en quelque sorte le père de la Déclaration de Paris, en raison des efforts qu'il a menés pour que les négociations aboutissent à un consensus sur un ensemble de bonnes pratiques de la gestion de l'aide. Le Forum de Paris a regroupé 500 représentants des pays donateurs et bénéficiaires de l'aide. De plus, de nombreux pays émergents ont été invités en tant qu'observateurs, dont la Chine, l'Inde et le Brésil. Toutefois, les acteurs non-membres du CAD-OCDE ont condamné la domination symbolique exercée par les hauts fonctionnaires occidentaux qui, selon leurs dires, s'est notamment exprimée par l'utilisation d'un vocabulaire technocratique. Les représentants du Sud qui n'étaient pas familiers des procédures de négociation au CAD se sont sentis entravés dans leur capacité à influencer les débats lorsqu'ils n'étaient pas réduits au statut d'observateurs[424]. Certains membres du CAD-OCDE eux-mêmes, par exemple la France ou le Japon, ont vu dans la Déclaration de Paris le fruit de la domination de la pensée anglo-saxonne sur le développement au sein de cet organisme[425].

Le CAD a répondu à ces critiques en incluant davantage de représentants des pays du Sud, de membres de la société civile internationale et d'organismes multilatéraux à la préparation du Forum de Haut Niveau sur l'efficacité de l'aide d'Accra en 2008. Le fait que le

---

[421](1) l'appropriation par les bénéficiaires de l'aide des programmes de développement, dont l'atteinte est évaluée par le critère d'existence de stratégies nationales de développement dans les pays partenaires ; (2) l'alignement des donateurs sur les besoins et capacités des pays partenaires évalué par l'existence de système de passation des marchés et de gestion des finances publiques fiables dans les pays partenaires, l'alignement des donateurs sur les priorités politiques nationales, le pourcentage de l'aide consacré au renforcement des capacités en évitant la mise en place de structures parallèles importées par les donateurs, l'utilisation des systèmes nationaux de gestion des finances publiques et de passation des marchés, une meilleure prévisibilité de l'aide et une aide non liée ; (3) l'harmonisation des politiques et pratiques des donateurs, qui s'évalue par l'utilisation de procédures ou de dispositifs communs et la mise en œuvre d'analyses conjointes ; (4) la gestion axée sur les résultats, dont l'atteinte se mesure par le nombre de pays dotés de cadres transparents d'évaluation des performances, se prêtant à un suivi régulier, et qui permettent d'évaluer les progrès par rapport aux objectifs fixés dans les stratégies nationales de développement et les programmes sectoriels ; (5) la responsabilité mutuelle dont la réalisation s'exprime par le nombre de pays partenaires où sont entreprises des évaluations mutuelles des progrès accomplis dans l'exécution des engagements souscrits concernant l'efficacité de l'aide, notamment ceux qui figurent dans la Déclaration de Paris.

[422] Pour une définition du terme « gestion axée sur les résultats », voir lexique en annexe 7.

[423] GOYETTE-COTE, Gabriel, *Le « nouveau régime de l'aide » à l'épreuve des faits : du Consensus de Washington à la Déclaration de Paris*, Mémoire de Science politique, Université du Québec à Montréal, mai 2010, 118p.

[424] Entretien avec un membre du réseau d'évaluateurs du CAD-OCDE 16 novembre 2012.

[425] Entretien avec un évaluateur de la JICA, 8 mai 2012 ; entretien avec un ancien directeur du département de l'évaluation de l'AFD, 21 décembre 2011.

forum ait lieu au Ghana est un symbole fort affichant une ouverture envers les pays du Sud. L'évènement rassembla environ un millier de délégués et une place plus active a été réservée aux États émergents qui ont contribué à la rédaction de certains documents finaux du forum, notamment sur la coopération Sud-Sud. De plus, ce forum a principalement porté sur la participation des bénéficiaires dans la gestion de l'aide, aspect que la société civile internationale avait jugé comme étant insuffisamment traité par la Déclaration de Paris[426]. Ces efforts n'ont toutefois pas permis d'enrayer les critiques.

Les Nations unies ont aussi travaillé, parallèlement à l'OCDE, à mieux intégrer les donateurs émergents et les pays bénéficiaires dans le régime international de l'aide, notamment avec l'organisation, en 2008, d'un Forum de la Coopération pour le Développement. Ce forum se voulait plus ouvert que les Forums de Haut Niveau du CAD-OCDE et avait comme principaux axes de travail la redevabilité mutuelle[427] et la coopération Sud-Sud. Nous voyons, encore une fois, apparaître la concurrence entre le CAD-OCDE et les Nations unies sur la production de bonnes pratiques sur l'efficacité de l'aide. En effet, la production de standards, normes, procédures, etc. permet aux organisations multilatérale d'élargir ou de renforcer leur mandat et de réaffirmer leur légitimité sur des enjeux donnés[428].

Cette tentative des Nations unies d'instaurer un dialogue alternatif avec les donateurs émergents et les bénéficiaires de l'aide a vite été sabordée par le CAD-OCDE qui, dès la création de ce forum, y a envoyé des représentants, afin de se poser en meneur des discussions[429]. Les chercheurs sont nombreux à souligner que ce forum est instrumentalisé par le CAD-OCDE comme une plateforme supplémentaire pour diffuser ses bonnes pratiques. Par exemple, Yash Tandon déclare : « les pays de l'OCDE ont efficacement saisi le Forum de la Coopération pour le Développement afin d'y importer leurs propres agendas et perspectives »[430]. Le CAD-OCDE a donc réussi à noyauter le Forum pour garder le monopole du dialogue sur l'efficacité du développement avec les acteurs du Sud et ainsi réaffirmer son mandat de plateforme d'échanges et de producteur de bonnes pratiques sur l'efficacité de l'aide au développement.

Avec le Haut forum d'Accra, le CAD-OCDE a significativement transformé son fonctionnement, dans la mesure où il a fait un effort significatif d'ouverture envers les États du

[426] DAC-OECD, *Programme d'action d'Accra*, 2008.
[427] Voir définition dans le lexique en annexe 7.
[428] KLEIN, Asmara, LAPORTE, Camille, SAIGET, Marie, 2015 (à paraître), *op. cit.*
[429] Entretien avec une ancienne Chief Social Development Advisor au DFID, consultante pour des agences de développement, 14 février 2014.
[430] TANDON, Yash, *Ending Aid Dependence*, Fahamu, 2nd revised edition, novembre 2008, 164p, pp. 123.

Sud et les acteurs non-étatiques du développement. Toutefois, leur intégration dans la structure technocratique qu'est l'OCDE ne fut pas plus aisée que lors du forum de Paris. Un délégué du Sud témoigne : « comme personne n'avait pensé à fournir aux représentants du Sud des explications sur le fonctionnement du CAD, ils durent, dans un premier temps, lutter pour suivre les discussions et comprendre leurs enjeux. Durant la « pause-café », un délégué commenta ironiquement qu'il était dommage que les participants du Sud n'aient été conviés qu'une fois le programme de travail du réseau établi, les empêchant ainsi d'influencer l'agenda politique de l'évènement »[431]. Nous voyons la domination symbolique forte qu'il peut y avoir dans les négociations internationales entre des acteurs habitués aux pratiques de négociations et au vocabulaire technocratique des organismes multilatéraux et des acteurs moins familiarisés à ces pratiques et à ce vocabulaire, se sentant vite dépassés. Cette domination s'atténue à mesure que les acteurs du Sud participent de plus en plus souvent et activement aux négociations et acquièrent ainsi des compétences nouvelles. En effet, les chercheurs ayant étudié le processus d'ouverture du CAD-OCDE notent une amélioration dans la participation des représentants du Sud : « les participants (très largement des hommes) sont capables de s'exprimer sur un ensemble varié d'enjeux politiques. Toutefois, leur présence dépend toujours du fait que les membres du CAD aient payé leurs billets d'avion et ils ne peuvent accéder aux négociations que sur invitation »[432]. Leur participation est donc encore largement dépendante du bon-vouloir des bailleurs occidentaux, ce qui peut limiter la liberté de leur prise de parole.

Malgré une inclusion de plus en plus large des acteurs divers du développement dans les forums internationaux du CAD-OCDE, ce sont bien les objectifs et les indicateurs de la Déclaration de Paris, négociés dans un cadre relativement fermé, qui ont prévalu dans les années 2000 comme référentiel international de l'efficacité de l'aide.

Contrairement aux OMD, les objectifs de la Déclaration de Paris ciblent explicitement les responsabilités des donateurs et des bénéficiaires[433]. Ce cadre normatif renforce ainsi la logique de performance et de concurrence dans laquelle sont placés les donateurs et les bénéficiaires, dans la mesure où leurs progrès sont mesurés et comparés. La mise en place d'indicateurs chiffrés et facilement communicables donne un pouvoir incitatif fort au CAD-OCDE envers les donateurs et les bénéficiaires qui vont chercher à se conformer aux bonnes

[431] EYBEN, Rosalind, 2013, *op. cit.*
[432] *Ibid*.
[433] Par exemple, en ce qui concerne l'alignement sur les programmes de développement nationaux, l'indicateur de réussite pour les donateurs est : « 50 % des apports de coopération technique sont mis en œuvre dans le cadre de programmes coordonnés conformes aux stratégies nationales de développement », et celui pour les bénéficiaires est : « la moitié des pays partenaires progressent d'au moins un degré (0.5 point) sur l'échelle de performance des systèmes de gestion des finances publiques/de l'évaluation de la politique et des institutions nationales », CAD-OCDE, 2005, *op. cit.*

pratiques de la Déclaration de Paris pour obtenir de meilleurs résultats que leurs pairs et apparaître comme des « bons élèves » sur la scène internationale.

Dans les discours officiels, la Déclaration de Paris est présentée comme une initiative visant à rendre plus équitable l'aide au développement, à donner plus de responsabilités aux pays partenaires et à rendre plus éthiques les pratiques des donateurs. Selon nous, ce discours cache une volonté de mieux contrôler et réguler la manière dont les États du Sud gèrent les fonds d'aide dans un contexte où l'aide budgétaire[434] croît.

La Déclaration de Paris a eu des conséquences importantes sur la pratique de l'évaluation. La « bonne pratique » de gestion axée sur les résultats requiert que les donateurs « relient la programmation et les ressources par pays aux résultats obtenus et alignent cette programmation et ces ressources sur les cadres d'évaluation des performances des pays partenaires. Les donateurs doivent s'abstenir de demander l'adoption d'indicateurs de performance qui ne soient pas en accord avec les stratégies nationales de développement »[435]. Les bénéficiaires, de leur côté, ont la responsabilité de « s'efforcer de mettre en place des cadres d'évaluation et de notification orientés vers les résultats, à l'aide desquels ils vont suivre les progrès réalisés dans la mise en œuvre des principaux volets des stratégies nationales et sectorielles de développement. Ces cadres devront permettre de suivre l'évolution d'un nombre raisonnable d'indicateurs pour lesquels il est possible de disposer de données à un coût abordable »[436]. Nous retrouvons dans ce principe de responsabilisation et dans ce vocabulaire très technocratique un marqueur du New Public Management, où l'évaluation est utilisée comme un outil managérial et de régulation des pratiques des acteurs.

D'autres bonnes pratiques ont aussi des conséquences sur la pratique et l'utilisation des évaluations. Les principes d'alignement[437] et de redevabilité mutuelle[438] appellent au renforcement des capacités en évaluation des pays du Sud pour qu'ils puissent évaluer les programmes d'aide au développement avec les donateurs. Le principe d'harmonisation requiert des donateurs qu'ils mènent des évaluations conjointes[439], afin de réduire les coûts des évaluations pour les donateurs et les bénéficiaires.

Nous voyons donc que l'amélioration de la pratique de l'évaluation est un élément clé de la réforme de la gestion de l'aide au développement portée par la Déclaration de Paris. En

[434] Pour une définition du concept « d'aide budgétaire », voir lexique en annexe 7.
[435] OECD, 2005, *op. cit.*
[436] *Ibid.*
[437] Pour une définition du concept d'alignement, voir lexique en annexe 7.
[438] Pour une définition du concept de redevabilité mutuelle, voir lexique en annexe 7.
[439] Pour une définition du concept d'évaluations conjointes, voir lexique en annexe 7.

filigrane de cette déclaration, nous voyons un idéal type de l'évaluation se dessiner, répondant aux critères suivants : des évaluations systématisées et orientées vers les résultats (gestion axée sur les résultats), participatives (renforcement des capacités et responsabilité mutuelle) et conjointes entre donateurs (harmonisation entre les donateurs). Une telle pratique de l'évaluation devrait mener à une diffusion de la culture de l'évaluation par les pays bénéficiaires, à une gestion plus efficaces de l'aide et à une meilleure appropriation des programmes d'aide par les pays bénéficiaires.

Dans la Déclaration de Paris, les donateurs s'engagent « à relier la programmation et les ressources par pays aux résultats obtenus et à aligner cette programmation et ces ressources sur les cadres d'évaluation des performances des pays partenaires »[440]. Ainsi, un nouvel usage des évaluations *ex post* devrait être de déterminer les montants d'aide à allouer aux pays en fonction de leurs résultats[441]. Pour atténuer l'image peu « humaniste » de cette sélectivité basée sur la performance, la déclaration précise que ces cadres d'évaluation des performances doivent être élaborés de manière collaborative avec les partenaires locaux, afin que ces derniers se fixent eux-mêmes les objectifs à partir desquels ils seront évalués.

En conclusion de cette partie, nous observons que les OMD présentent l'évaluation comme un outil permettant de rendre l'aide plus équitable et efficace dans la lutte contre la pauvreté. L'agenda sur l'efficacité de l'aide met, quant à lui, en avant une utilisation gestionnaire de l'évaluation qui devient un outil de rationalisation de la mise en œuvre de l'aide et de dépolitisation des relations entre donateurs et bénéficiaires. Ces deux approches de l'évaluation apparaissent, sur certains points, contradictoires, d'autant plus qu'elles sont défendues par les mêmes acteurs qui oscillent entre l'idéal d'une aide universalisée et égalitaire et la réalité d'une aide sélective et qui donne une place centrale à l'obtention de résultats à court terme. Au final, c'est l'approche gestionnaire de l'évaluation qui domine dans la pratique des organismes d'aide, en quête de preuves de résultats, de légitimité et de redevabilité. La montée en importance de cette approche de l'évaluation n'est pas seulement due aux transformations qui touchent la politique internationale du développement ; comme nous allons le démontrer, elle tire également ses origines de tendances observables au sein des politiques publiques nationales des États donateurs.

[440] OECD, 2005, *op. cit.*
[441] Cet usage de l'évaluation sera davantage étudié dans le chapitre 8 de la thèse.

B. Une demande politique croissante d'évaluations expérimentales

La complexification des méthodes évaluatives est l'une des évolutions qui ont le plus marqué le secteur de l'évaluation de l'aide au développement dans les années 2000. Ce phénomène s'observe notamment par la montée en puissance de la méthode expérimentale d'évaluation des impacts de l'aide[442]. Cette méthode, d'abord prônée par un petit groupe de chercheurs et de bailleurs de fonds, a rapidement acquis le statut de « standard d'or » de l'évaluation. Nous allons désormais étudier les causes de la popularité exponentielle de la méthode expérimentale et ce qu'elle reflète de la pensée politique actuelle de l'aide au développement.

- **L'application des méthodes expérimentales à l'aide au développement**

Les chercheurs du *The Abdul Latif Jameel Poverty Action Lab* (J-PAL), laboratoire américain de recherche sur la lutte contre la pauvreté, sont les premiers à avoir théorisé et appliqué une méthode « aléatoire expérimentale » à l'évaluation des projets d'aide au développement. Empruntée au domaine médical, cette méthode consiste à « sélectionner aléatoirement dans une population des groupes d'individus qui participeront au programme évalué, de telle sorte que le groupe d'individus traité soit statistiquement similaire au groupe d'individus non traité. Les différences observées dans chacun des groupes sont alors interprétées comme l'effet causal de l'intervention »[443].

La montée en puissance des évaluations expérimentales a été permise par le soutien financier de grandes fondations anglo-américaines (la *Bill et Melinda Gates Fundation*, la *William and Flora Hewett Fundation*) et d'agences d'aide (la Banque mondiale, le DFID, ou encore USAID). À ce sujet, Rosalind Eyben raconte : « le directeur exécutif du J-PAL et le Président du *Center for Global Development* (CGDev) créditent la Fondation Gates d'avoir été l'un des principaux défenseurs de *l'evidence based approach* du développement, influençant la manière dont les gouvernements allouaient leurs financements »[444].

En plus de financer les recherches du J-PAL, ces acteurs ont, par exemple, soutenu *l'Initiative Internationale pour les Évaluations d'Impact* (3IE), lancée en 2008 par le CGDev,

[442] En anglais : *randomized control trials.*
[443] LABROUSSE, Agnès, « Nouvelle économie du développement et essais cliniques randomisés : une mise en perspective d'un outil de preuve et de gouvernement », *Revue de la régulation,* n°7, 2010.
[444] EYBEN, Rosalind, 2013, *art. cit.*

à la suite du constat du déficit de connaissances rigoureuses sur la mesure de l'impact des programmes de développement sur la réduction de la pauvreté. Les objectifs officiels du 3IE sont : de produire et diffuser des preuves de ce qui est efficace dans l'aide au développement, de construire une culture de *l'evidenced-based policy* et d'encourager la mise en œuvre et l'utilisation des évaluations d'impact. Une autre fonction officieuse est la création d'une communauté épistémique de chercheurs sur les évaluations d'impact, afin d'augmenter leur visibilité et leurs poids et de promouvoir plus efficacement leur expertise[445].

C'est également grâce au succès de l'évaluation expérimentale du programme *Progresa* de microcrédit au Mexique, réalisée par *l'International Food Policy Institute*, que la méthodologie expérimentale d'évaluation a connu un grand succès. Le programme consistait à effectuer des virements bancaires à des familles rurales pauvres, à condition qu'elles participent à des programmes de santé et de nutrition et qu'elles envoient leurs enfants à l'école. Ce programme a été évalué selon une méthode expérimentale aléatoire contrôlée avec la constitution de groupes bénéficiaires et de groupes « témoins » non bénéficiaires. L'histoire veut que les preuves rigoureuses de l'efficacité du programme apportées par l'évaluation expérimentale auraient persuadé le gouvernement d'élargir ce dispositif. Cette histoire a été relayée par les réseaux sociaux, des ONG et des donateurs. En conséquence, elle a conduit à la mise en œuvre d'autres programmes de ce type dans le monde, dont la plupart ont également été évalués avec une méthode expérimentale[446]. Selon nous, l'utilisation politique directe qui a été faite de l'évaluation par le gouvernement mexicain a largement contribué à populariser cette méthode dans le secteur de l'aide. Elle répondait à la volonté des dirigeants des agences d'aide d'avoir des évaluations qui aient une influence directe sur la prise de décision politique.

La popularisation de la méthodologie des expérimentations aléatoires s'exprime aussi dans l'intérêt qu'y consacrent les universités les plus prestigieuses (MIT, Harvard, Yale, etc.) et par la multiplication des articles scientifiques référencés utilisant cette méthode. Agnès Labrousse, dans un article critique sur la méthode expérimentale, observe que « la moitié des articles publiés dans le *Quarterly Journal of Economics* en 2008 utilisaient des données microéconomiques de pays en développement correspondant à des expériences randomisées »[447]. Les organisations multilatérales ont également développé leur propre expertise en évaluation expérimentale. La Banque mondiale, par exemple, a lancé la

[445] Site internet de l'initiative 3IEhttp://www.3ieimpact.org/ consulté le 13 novembre 2014.
[446] BORUCH, Robert, *Place randomized Trials: Experimental Test of Public Policy*, The annals of the American Academy of Political and Social Science, Philadelphia, 2005, 286p, pp. 208-217; SKOUFIAS, Emmanuel, *Progresa and its Impacts on the Welfare of Rural Households in Mexico*, Research Report n°139, International Food Policy Research Institute, 2005.
[447] LABROUSSE, Agnès, 2010, *art. cit.*

Development impact evaluation initiative[448] qui a pour objectif de promouvoir l'utilisation des méthodes expérimentales dans l'évaluation des programmes d'aide. Face à cet engouement généralisé, les plus gros donateurs bilatéraux, notamment les États-Unis, le Royaume-Uni et la France, ont mobilisé des fonds et de l'expertise pour mettre en œuvre des évaluations expérimentales[449]. Nous observons donc ici un effet d'émulation internationale.

Le succès de la méthode expérimentale a été tel que la signification du terme « évaluation d'impact » en a été modifiée. Jusqu'alors, le CAD-OCDE définissait l'évaluation d'impact comme « l'évaluation des changements positifs et négatifs produits par une intervention de développement, directement ou indirectement, de manière voulue ou non voulue »[450]. Nous avons vu dans les premiers chapitres que les évaluations d'impact étaient réalisées selon des méthodes qualitatives et participatives. Aujourd'hui, l'évaluation d'impact tend à être systématiquement associée à la mesure expérimentale des effets des programmes d'aide sur le long terme. Par exemple, la politique d'évaluation d'USAID définit l'évaluation d'impact comme « les comparaisons faites entre des bénéficiaires auxquels a été assignée aléatoirement une intervention de développement et un groupe de contrôle d'individus non-bénéficiaires aux caractéristiques identiques »[451] .

Nous pensons que la méthode évaluative expérimentale attire les bailleurs de fonds et leurs tutelles financières car elle donne une image dépolitisée et scientifique à l'évaluation et, par extension, à l'efficacité de l'aide. Elle permettrait un usage direct des évaluations par les décideurs politiques. Elle redonne à l'évaluation une fonction d'expression d'une vérité sur ce qui fonctionne et ne fonctionne pas dans l'aide au développement, fonction qu'elle avait perdue depuis les années 1960. Au contraire, nous allons démontrer que les tenants de cette méthodologie, revendiquant faire des analyses scientifiques et neutres des projets d'aide, véhiculent des convictions politiques sur la nature et le rôle que devraient avoir l'aide au développement. De manière générale, nous constatons que la popularisation de nouvelles méthodes d'évaluation et les débats qu'elles entraînent sont toujours le reflet d'évolutions plus profondes des fondements de l'aide au développement. C'est ce lien entre une évolution des

[448] Site de l'initiative : http://web.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTDEVIMPEVAINI/0,,contentMDK:23422972~pagePK:64168445~piPK:64168309~theSitePK:3998212,00.html , consulté le 15 novembre 2014.

[449] En 2014, l'équipe d'évaluation de l'AFD compte un spécialiste des méthodes expérimentales et avait six évaluations expérimentales en cours ou achevés ; le DFID et USAID développent une recherche importante sur le sujet et organisent de nombreuses journées d'étude sur le thème.

[450] CAD-OCDE, *Glossaire du CAD de l'OCDE sur les principaux termes relatifs à l'évaluation*, 1991.

[451] USAID, *Evaluation: Learning from Experience,* USAID Evaluation Policy, janvier 2011.

pratiques d'évaluation et des fondements politiques de l'aide que nous allons, encore une fois, tenter d'expliciter.

Les évaluateurs utilisant des méthodes expérimentales se prévalent de pouvoir analyser au niveau « micro » les critères d'efficacité des projets en faisant une comparaison entre l'évolution des conditions de vie des populations qui bénéficient d'un projet (groupe traité) d'aide et celles qui n'en bénéficient pas (groupe de contrôle). Jusqu'alors, l'efficacité de l'aide était majoritairement évaluée par la méthode macroéconomique des « régressions en coupe internationale » qui cherchait les facteurs communs de croissance entre plusieurs pays recevant le même type de projets d'aide. Cette méthode permettait seulement d'émettre des liens de corrélation entre un certain type d'aide au développement et des progrès de développement observés. Il était impossible d'isoler la contribution de l'aide des autres facteurs de développement structurels et conjoncturels propres à chaque pays. Les donateurs ont donc vu dans les évaluations expérimentales un bon outil de redevabilité, car elles permettent d'isoler les effets spécifiques de leurs interventions sur lesquels ils peuvent ensuite rendre compte.

La critique de la méthode des « régressions en coupe internationale » est un argument fort des tenants de la méthode des évaluations expérimentales, notamment Esther Duflo et Abhijit Banerjee, pour faire valoir les avantages de la méthode qu'ils ont théorisée. Ils tiennent des discours du type : « même avec deux millions de régressions, nous n'arriverons pas à percer le secret de la croissance »[452]. Selon eux, les évaluations d'impact aléatoires contrôlées constitueraient une avancée majeure, dans la mesure où elles apporteraient des « preuves scientifiques » sur l'efficacité de l'aide et permettraient d'isoler les effets de l'aide sur les populations les plus vulnérables. Elles seraient donc l'instrument idéal pour vérifier que l'aide bénéficie bien aux plus pauvres. Cette méthode s'inscrit donc parfaitement dans un contexte où l'aide « s'individualise », se concentre sur la lutte contre la pauvreté, et requiert des preuves de résultats « scientifiques ».

Nous retrouvons chez les tenants des expérimentations aléatoires l'idée selon laquelle l'évaluation peut avoir une influence directe sur la prise de décision politique. La théorie politique de l'acteur rationnel[453] est très présente dans leurs discours. Esther Duflo déclare par exemple : « il s'agit de rendre les résultats (des évaluations) présents dans le discours politique. Il n'y a certes pas que les résultats qui déterminent l'adoption d'une politique, mais ça joue

[452] DUFLO, Esther, *Expérience, science et lutte contre la pauvreté*, Fayard, 2009, 60p, pp. 23-24.
[453] La théorie de l'acteur rationnel a notamment été développée par George Homans dans l'ouvrage *Social Behavior: Its Elementary Forms* publié en 1974, ou par Robert Axelrod dans son étude de la théorie des jeux : T*he Evolution of Cooperation*, Basic Books, 1984, 264p.

quand même un rôle important »[454]. Les tenants de cette approche ont l'ambition de transformer les fondements politiques de l'aide. Ils pensent que « créer une culture promouvant, encourageant et finançant des évaluations rigoureuses par expérimentations randomisées, peut potentiellement révolutionner les politiques sociales du 21e siècle, comme les essais randomisés avaient révolutionné la médecine du 20e siècle »[455]. Sous couvert de l'apport d'une méthodologie neutre et scientifique, apparaît donc bien une pensée politique qui entend refonder les principes de l'aide au développement.

La méthode expérimentale d'évaluation d'impact est de plus en plus sujette à la critique. Critique que nous rejoignons en partie. Les premiers détracteurs de ces évaluations ont été des chercheurs en économie qui en contestaient la valeur ajoutée en termes d'analyse de l'efficacité de l'aide. Angus Deaton[456], chercheur en économie à l'université de Princeton et Dani Rodrik[457], chercheur en économie à l'université de Harvard, estiment, par exemple, que les résultats des évaluations d'impact aléatoires contrôlées ne sont pas généralisables, car ils ne sont valables que sur un échantillon non représentatif de la population d'une région ou d'un pays dans un contexte isolé. Selon eux, ces évaluations sont un outil remarquable d'isolation de l'effet d'une action spécifique de développement, mais elles ne saisissent pas les interactions entre les différentes actions de développement et entre le projet d'aide et le contexte politique, économique et social du pays. Or, selon nous, le développement d'un groupe d'individus n'est viable que dans un contexte plus large de développement. Il est le résultat complexe d'une chaîne d'effets d'entrainement entre les niveaux international, national et local de développement. Les évaluations expérimentales ne résolvent donc pas le problème de la difficile mesure de l'efficacité « macro » de l'aide et ne peuvent pas juger de la durabilité et de la reproductibilité des projets d'aide.

Selon Agnès Labrousse, une autre faiblesse de la méthode des expérimentations aléatoires est « qu'elles se concentrent sur les résultats des projets d'aide sans analyser les mécanismes causaux qui ont mené à ces résultats »[458]. Nous expliquons cela par le fait que l'approche expérimentale se fonde sur un présupposé de rationalité des acteurs. Les mécanismes causaux sont pensés comme étant automatiques et déterminables *a priori* et ne nécessitant pas

[454] DUFLO, Esther, in MANERY, Florian, « L'économie du développement à l'épreuve du terrain, Entretien avec Esther Duflo », *la vie des idées*, 5 mai 2009.
[455] DUFLO, Esther, in DEATON, Angus, *Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development*, Manuscript, Princeton, janvier, 2009.
[456] DEATON, Angus, *Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development",* Manuscript, Princeton, janvier 2009.
[457] RODRIK, Dani, *The new development economics: we shall experiment, but how shall we learn?* Paper prepared for the Brookings Development Conference, 29-30 mai 2008.
[458] LABROUSSE, 2011, *art. cit.*

d'explication particulière. Ces évaluations n'apportent donc pas d'informations sur les processus sociaux ou politiques qui ont permis l'atteinte de tel ou tel résultat, ce qui remet en cause leur utilité en terme d'apprentissage et d'amélioration de l'efficacité de l'aide. La méthode semble peu adaptée pour comprendre, par exemple, pourquoi des campagnes de sensibilisation sur le sida ont des impacts positifs, neutres ou négatifs, résultats qui sont largement déterminés par le contexte social, politique et culturel de l'action.

Par ailleurs, après avoir lu une vingtaine de rapports d'évaluations expérimentales, nous constatons qu'elles ne remplissent pas leur objectif d'évaluer l'impact de l'aide au développement sur la réduction de la pauvreté des pays bénéficiaires. Elles ne parviennent souvent qu'à évaluer les résultats intermédiaires des projets. L'exigence d'apporter des preuves quantifiées et objectives incite les évaluateurs à se concentrer sur les réalisations des projets plus faciles à mesurer rigoureusement que leurs résultats ou leurs impacts. Par exemple, sur un projet agricole, les évaluateurs regardent si les agriculteurs utilisent davantage d'engrais chimiques (objectif intermédiaire), mais ne vont pas jusqu'à s'intéresser aux effets du projet sur la réduction de la pauvreté de ces agriculteurs (objectif final)[459]. Rosalind Eyben estime, quant à elle, que « *la hausse de l'utilisation des méthodes expérimentales d'évaluation et de la culture du résultat dans le secteur de l'aide a entraîné une éviction des activités dont les résultats sont les moins mesurables (le renforcement des capacités institutionnelles) au profit de celles dont les résultats sont les plus mesurables (la délivrance de biens et services)* »[460]. Nous observons ici comment l'évaluation peut avoir un effet indirect sur les pratiques de développement.

L'effet de miroir entre les changements politiques de l'aide au développement dans les années 2000 (individualisation, équité, gestion axée sur les résultats, redevabilité accrue) et le discours entourant la méthodologie des expérimentations aléatoires est tellement fort que ce dernier en reflète même les paradoxes. Nous remarquons, en effet, que les déclarations d'Esther Duflo reproduisent l'ambiguïté entre une approche rationnelle et quantitative du développement et une inflexion qualitative où les processus sociaux ont plus de place : « je souhaite pratiquer l'économie comme une vraie science humaine. Une science rigoureuse et impartiale. Une science de l'homme, dans toute son imperfection et sa complexité. Une science humaine :

[459] LABROUSSE, 2011, *art, cit.*
[460] Entretien avec une ancienne Chief Social Development Advisor au DFID, consultante pour des agences de développement, 14 février 2014.

humble et condamnée à l'erreur »[461]. Ici encore, la manière dont le lien peut être fait entre ces deux approches, « une science impartiale » et une « entreprise humaine », n'est pas explicitée.

Nous voyons donc que la montée en puissance de l'évaluation expérimentale s'est faite en réponse à un contexte et une demande politiques particuliers.

- **Évaluation scientifique et gouvernement par la performance**

Nous allons démontrer, qu'au-delà du contexte politique international de l'aide au développement, la popularisation de l'évaluation expérimentale répond à une demande politique nationale dans les pays donateurs. Il nous semble, en effet, réducteur de considérer que seule l'influence normalisatrice des organisations multilatérales suffit à transformer les pratiques des bailleurs de fonds bilatéraux. Le relai politique au niveau national est fondamental, car les agences d'aide bilatérales sont avant tout redevables devant les décideurs politiques nationaux, avant de l'être auprès de leurs partenaires internationaux. Si les agences d'aide se sont mises à développer une expertise en matière d'évaluation expérimentale, c'est donc parce qu'elles ont été portées par un contexte politique national allant dans ce sens.

En ce qui concerne le Royaume-Uni, l'évaluation des politiques publiques a connu une institutionnalisation et un essor important dès 1997 avec l'arrivée de Tony Blair au pouvoir. Rosalind Eyben explique ce phénomène : « au Royaume-Uni, en 1997, le gouvernement Labour était très enthousiaste à l'idée de démontrer que ses décisions ne seraient pas guidées par une idéologie socialiste, mais par des preuves non-biaisées et techniques »[462]. Nous identifions ici les débuts de la montée de l'*evidence based policy* au Royaume-Uni qui a été à l'origine de la demande pour les évaluations expérimentales, ces dernières permettant d'apporter les preuves sur lesquelles fonder la prise de décision.

Cette tendance s'est confirmée dans les années 2000, marquées notamment par une volonté du gouvernement d'harmoniser les pratiques d'évaluation des administrations. C'est ainsi que le *Magenta Book*, publié par le gouvernement en 2003, émet des recommandations sur les méthodes à utiliser pour mener une évaluation, sur les forces et faiblesses des différents types d'évaluations et sur le recueil des données nécessaires à l'analyse des politiques publiques. Par ailleurs, plusieurs autres ressources sur la politique d'évaluation du gouvernement ont été produites pour aider les administrations à comprendre l'utilité des

[461] DUFLO, 2009, *op. cit.* p. 75.
[462] EYBEN, 2013, *art. cit.*

évaluations et à les réaliser de manière homogène et cohérente[463]. Selon Colin Talbot, professeur de science politique à l'université de Manchester : « bien qu'il n'existe pas de données sur les dépenses totales et le niveau d'activité de l'évaluation des politiques publiques (au Royaume-Uni), la vision générale (d'après des entretiens avec des acteurs clés) est que l'activité évaluative s'est énormément étendue dans les années 2000 »[464]. En effet, la base de données gouvernementale qui recense les rapports d'évaluations de l'administration en compte un nombre important (entre 100 et 150 tous les six mois)[465], auquel s'ajoutent les rapports d'évaluation externes des lobbies, groupes d'intérêts, think-tanks, etc.

La modernisation de l'État est assimilée à l'usage de pratiques expérimentales d'évaluation, vues comme un outil de rationalisation de la prise de décision et des dépenses budgétaires. Cette approche est décrite dans le *Modernising Government agenda*[466] en 1999 et dans le rapport *Measuring the Performance of Government Departments*[467] en 2001. Selon Colin Talbot, l'influence des « preuves » obtenues par les évaluations sur la pratique des politiques publiques semble assez faible, malgré les ressources supplémentaires allouées pour leur mise en œuvre. Toutefois, il souligne « la difficulté d'estimer cette influence en l'absence de mécanismes de contrôle sur l'utilisation ou non des résultats des évaluations dans la prise de décision politique »[468]. Nous en concluons que si les politiques sont rendues plus transparentes par les évaluations, les leçons qui sont tirées ou non de ces dernières restent une inconnue.

Nous constatons donc qu'au Royaume-Uni la demande politique d'évaluations expérimentales s'est développée en même temps que, voire a précédé, la diffusion de ce type d'évaluations dans les instances multilatérales, le pays a toujours eu une culture forte de la quantification et de l'empirisme dans la pratique et l'analyse des politiques publiques. C'est un trait culturel que les États anglo-américains ont su diffuser au sein des organisations multilatérales de développement et qui a ensuite imprégné les politiques publiques d'autres États comme la France.

La politique d'aide au développement du Royaume-Uni est clairement gagnée par la hausse d'une demande en évaluation et de transparence depuis la fin des années 2000. Dans son

[463] HM Treasury, *the « Green Book »: appraisal and evaluation in Central Government: Treasury guidance,* 2003; *The Better Regulation Executive,* The Department of Work and Pensions, *Research Methods on Policy Evaluation,* 2001.
[464] TALBOT, Colin, *Performance in Government : The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom,* Evaluation Capacity Development, IEG World Bank, novembre 2010.
[465] Les rapports d'évaluation sont consultables sur le site internet du gouvernement britannique https://www.gov.uk/government/collections/evaluation-reports, consulté le 6 décembre 2014.
[466] Prime Minister and the Minister for the Cabinet Office, *Modernising Government,* mars 1999.
[467] NAO, *Measuring the Performance of Government Departments,* 22 mars 2000.
[468] *Ibid.*

discours d'investiture au Secrétariat d'État à la coopération internationale, en 2010, Andrew Mitchell, directeur du DFID, déclarait : « ma priorité première est de sécuriser la *value for money* maximale de l'aide, à travers une meilleure transparence, des évaluations rigoureuses et indépendantes et un effort continu en matière d'obtention de résultats »[469]. Cette stratégie a été louée par le CAD-OCDE quelques années plus tard : « le programme d'aide du Royaume-Uni est de plus en plus soumis à des exigences de redevabilité, à mesure que son budget augmente (et en raison de la crise économique). Le DFID a fait des efforts pour démontrer son efficacité en liant ses actions aux résultats et en augmentant son travail de communication »[470]. Pour institutionnaliser ces bonnes pratiques, le gouvernement britannique a adopté en 2006 *l'International Development Reporting and Transparency Act* qui oblige le DFID à rendre compte, sous la forme de rapports annuels d'évaluation, des résultats de la politique d'aide britannique[471].

En France, selon Pierre Lascoumes, chercheur en sociologie politique à Sciences Po Paris et Clément Lacouette-Fougère, consultant en évaluation des politiques publiques, la réforme de la LOLF adoptée en 2001 « a conduit à transformer la vision des politiques publiques en une série de missions (composées de programmes et d'actions) axées autour d'objectifs et devant produire des résultats mesurés à l'aide d'un petit nombre d'indicateurs » [472]. Dans la lignée de cette réforme, la Révision générale des politiques publiques (RGPP) a été lancée en 2005, afin de « faire examiner les objectifs, les dépenses, les résultats et les *modus operandi* des grandes politiques publiques ministérielles par des équipes d'audit »[473]. Nous identifions là une volonté affichée d'externaliser le contrôle des politiques publiques dans une logique d'indépendance et de redevabilité. Nous retrouvons l'esprit de la gestion axée sur les résultats développé dans la Déclaration de Paris. Enfin, en 2013, dans le cadre de la Modernisation de l'action publique (MAP), le Secrétariat général pour la modernisation de l'action publique fut créé, dont l'un des objectifs est de mettre en place une collaboration interministérielle et un échange de bonnes pratiques en matière d'évaluation. Dans cette initiative, nous voyons une volonté de sortir d'un modèle de diffusion *top down* de la culture de l'évaluation, qui n'a jamais permis une réelle appropriation des méthodes évaluatives par les administrations, pour envisager une diffusion horizontale basée sur l'échange d'expériences entre les administrations.

---

[469] http://www.dfid.gov.uk/news/latest-news/2010/andrew-mitchell-appointed-secretary-of-state/, consulté le 6 décembre 2014.
[470] DAC-OECD, *Peer Review: United Kingdom*, Paris, 2010.
[471] *International Development (Reporting and Transparency) Act*, 2006.
[472] LACOUETTE-FOUGERE, Clément, LASCOUMES, Pierre, *Les scènes multiples de l'évaluation : les problèmes récurrents de son institutionnalisation*, Sciences-po, LIEPP, mai 2013, 122p, p. 63.
[473] *Ibid*.

En France, les réformes de la gestion des politiques publiques entretiennent une confusion commune entre le contrôle, l'audit et l'évaluation[474]. En effet, aussi bien dans la LOLF que dans la RGPP, si le vocabulaire de l'évaluation est utilisé et les références aux critères de qualité du CAD-OCDE fréquentes, la pratique évaluative se confond avec celles de l'audit et du contrôle et reste majoritairement cantonnée à des aspects d'efficience des politiques publiques, en laissant les enjeux sociaux et les résultats à la marge. L'évaluation risque en conséquence de perdre sa spécificité par rapport au contrôle et à l'audit[475]. Pour illustrer ce propos, l'exemple de la Cour des comptes[476], qui développe depuis quelques années une fonction d'évaluation des politiques publiques, montre les difficultés qu'éprouvent les hauts fonctionnaires à comprendre la valeur ajoutée de l'évaluation par rapport au contrôle qu'ils pratiquent traditionnellement. Par exemple, certains fonctionnaires de la Cour ne parviennent pas à concevoir l'utilité d'intégrer aux études une analyse de la satisfaction des bénéficiaires des politiques publiques *via* des méthodes d'enquête qualitatives (questionnaires, ateliers de travail avec des groupes de bénéficiaires) et spécifiques à l'évaluation (cadre logique, sociogramme[477], etc.) jugées peu rigoureuses et technocratiques. Ainsi, les rapports d'évaluation produits par la Cour ne sont aujourd'hui qu'une façade. Ils reprennent les standards internationaux du CAD-OCDE et, en partie, le vocabulaire de l'évaluation (redevabilité, gestion axée sur les résultats, cadre logique, etc.), mais se fondent sur une méthodologie largement quantitative et axée sur le contrôle des procédures administratives des organisations, plutôt que sur des méthodes qualitatives explorant les processus de mise en œuvre, les résultats et l'utilité des politiques publiques. Ces rapports sont donc des « hybrides » entre contrôle et évaluation, qui témoignent, au mieux, de l'appropriation en cours de la fonction d'évaluation dans cet organisme public. En France, la demande politique en évaluation est donc peu éclairée sur la spécificité de l'évaluation par rapport au contrôle et à l'audit.

À ces vagues de réformes inspirées de l'idéologie anglo-américaine du New Public Management, s'est ajoutée, dans les années 2000, la montée en puissance du courant de *l'evidence based policy* qui pousse encore plus loin le besoin de résultats et de preuves quantifiables dans la gestion des politiques publiques. Alors que le New Public Management renvoie, comme son nom l'indique, à une vision gestionnaire de la performance qui se concentre

[474] La distinction entre ces trois disciplines est expliquée en introduction et dans le lexique.
[475] Pour une explication de la distinction faite entre évaluation, audit et contrôle, se référer à l'introduction ou au lexique.
[476] Ces observations sont tirées d'une mission d'expertise menée au sein de la Cour des comptes de mars à juin 2013, dans l'objectif d'apporter un appui méthodologique aux membres de l'organisme pour la réalisation d'une évaluation sur le réseau culturel français à l'étranger.
[477] Schéma représentant les différentes parties prenantes d'une politique publique et leur statut (financeurs, opérateurs, bénéficiaires) au sein de cette politique publique.

sur les activités et les outputs des programmes d'aide, *l'evidence based policy* est une notion développée par le milieu de la recherche qui se penche sur les impacts des programmes d'aide. Selon ce courant de pensée, les impacts des politiques publiques devraient être prouvés de manière scientifique et objective, ces preuves servant ensuite la prise de décision. D'après un évaluateur de l'AFD spécialisé dans les méthodes expérimentales, « *il s'agit d'un plaidoyer pro domo des chercheurs (économistes) pour qu'on fonde la décision publique sur de la bonne science* »[478].

Dans un contexte où se mêlent New Public Management et *evidence based policy*, les responsables politiques français ont développé une demande d'évaluations expérimentales aléatoires à la fin des années 2000. Le premier à populariser cette démarche, en 2007, a été Martin Hirsch, alors Haut-Commissaire à la jeunesse, au sujet de la politique du revenu de solidarité active. Il a soutenu la mise en place d'une évaluation expérimentale aléatoire en amont de la mise en œuvre de la politique, afin de mettre en place le protocole scientifique adapté[479]. Même si cette première initiative d'évaluation expérimentale à grande échelle n'a finalement pas pu être menée de manière rigoureuse, étant donné que les groupes bénéficiaires et de contrôle n'ont pas été choisis de manière aléatoire[480], la méthode s'est ensuite diffusée en tant que pratique d'excellence en matière d'évaluation. Depuis cette première étude, des dizaines de programmes sociaux innovants font l'objet d'évaluations aléatoires chaque année[481].

Cette diffusion de la méthodologie expérimentale aléatoire est aussi due à la montée d'une offre d'expertise. Le retentissement des travaux d'Esther Duflo[482], et plus largement, de l'École d'Économie de Paris qui se fait le chantre des évaluations expérimentales, a eu une influence cruciale dans la popularisation de cette méthode. L'argument principal d'Esther Duflo pour promouvoir les méthodes expérimentales dans le champ de l'évaluation des politiques publiques est « puisque l'évaluation expérimentale a fait ses preuves pour analyser les causes

[478] Entretien avec un évaluateur de l'AFD spécialisé dans les méthodes expérimentales, 19 novembre 2013.
[479] L'HORTY, Yannick, PETIT, Pascale, « L'évaluation aléatoire : un succès qui ne doit rien au hasard », *La vie des idées,* 11 mai 2011.
[480] Dans le cadre de l'expérimentation du RSA, ni la liste des départements expérimentateurs, ni le périmètre des zones tests dans chaque département, ni la liste des allocataires du RSA n'ont été choisi au hasard. Les départements, qui sont les véritables pilotes du RMI depuis la loi de décentralisation de décembre 2003, ont défini le périmètre des zones test sur la base d'une sélection raisonnée, en appliquant des critères et selon des contraintes qui leurs sont propres. À l'intérieur de ces zones, tous les allocataires du RMI bénéficient du RSA.
L'expérimentation du RSA s'inscrit donc dans le registre des quasi-expériences. Chaque département réalise une expérience qui est au niveau national répétée plus de trente fois, selon des modalités qui varient à la marge. On est bien dans la situation la plus courante de l'évaluation des politiques publiques où les bénéficiaires de la politique ne font pas l'objet d'un tirage au sort. La définition des zones témoins a été adaptée en conséquence (Goujard et L'Horty, 2010).
[481] *Ibid.*
[482] Professeure au MIT, membre du réseau international J-PAL et professeure au Collège de France.

de la pauvreté dans les pays pauvres, il s'agit maintenant de l'utiliser pour le même objectif dans les pays riches et notamment en France »[483]. Cette déclaration montre le caractère orignal de la diffusion de cette approche de l'évaluation dans l'espace international. Elle s'est, à l'origine, développée aux États-Unis dans les années 1960 et sa diffusion est alors restée cantonnée aux pays anglo-saxons. C'est après un transfert de cette méthodologie au niveau multilatéral, dans le champ de l'aide au développement, qu'elle s'est diffusée à l'ensemble des pays occidentaux, grâce à la constitution de réseaux de recherche internationaux sur cette thématique dont les chercheurs ont promu la méthode dans leurs pays d'origine. C'est le consensus entre un ensemble de chercheurs, de décideurs politiques et de praticiens des politiques publiques qui a érigé l'évaluation expérimentale au statut de bonne pratique de la gestion des politiques publiques. Nous voyons ainsi de quelle manière les bonnes pratiques de politiques publiques peuvent circuler dans l'espace international, par des boucles de rétroaction, entre les niveaux national et multilatéral. Par ailleurs, il est assez paradoxal de voir une méthodologie aussi technocratique que celle des évaluations expérimentales connaître un tel succès dans un pays comme la France, où la culture de l'évaluation est encore peu présente et où elle est souvent qualifiée de trop technique et déshumanisée.

À côté de ces discours et réformes politiques favorables à une évaluation quantifiée de la performance de l'État, où l'évaluation est considérée comme un outil de rationalisation budgétaire et de mesure du ratio coût/performance des politiques publiques, se développe un plaidoyer en faveur d'une approche participative de l'évaluation, qui inclurait les citoyens et permettrait de rendre plus redevable et ouverte la gestion des politiques publiques. De nombreux chercheurs théorisent des approches participatives de l'évaluation, afin de rompre avec une pratique de l'évaluation « par et pour les administrations »[484]. Toutefois, nous avons constaté, au cours d'observations participantes à la Société française d'évaluation et de stages dans des cabinets d'évaluation en 2011, que les tenants des deux approches n'échangent guère et qu'elles se développent de manière autonome. L'approche participative de l'évaluation se développe surtout dans les collectivités territoriales, notamment dans le secteur du développement durable ou des politiques d'urbanisation. Au niveau des politiques nationales, c'est avant tout une approche quantitative de l'évaluation qui prévaut et qui s'institutionnalise.

[483] DUFLO, Esther, 2009, *op. cit.* p. 42.

[484] Voir par exemple : BARON, Gaëlle, MONNIER, Éric, « une approche pluraliste et participative: coproduire l'évaluation avec la société civile », *Revue informations sociales*, n°110, septembre 2003, pp. 1-21 ; FAURE, Alain, SMITH, Andy, « l'évaluation objet de savoir et enjeu de pouvoir », *Sciences de la société*, n°32, pp. 149-152 ; THOENIG, Jean-Claude, « l'évaluation source de connaissance applicable aux réformes de la gestion politique », *Revue française d'administration publique*, n°92, 1999.

C. Des « bonnes pratiques » de l'évaluation difficiles à opérationnaliser

Nous allons désormais étudier comment les standards internationaux d'évaluation ont été appropriés dans les organismes d'aide. Nous démontrerons que si l'évaluation se systématise, les évaluateurs peinent à se conformer aux bonnes pratiques prônées par le CAD-OCDE.

- **Une pratique systématisée de l'évaluation**

Nous avons vu que le CAD-OCDE formule les standards de l'évaluation qui prévalent internationalement depuis le début des années 1990. À mesure que leur application s'est répandue, la qualité des évaluations des agences d'aide a elle-même été soumise à des évaluations, soit à la demande du CAD-OCDE, soit des bailleurs de fonds eux-mêmes, qui souhaitent avoir du recul sur leurs activités d'évaluation. Nous rentrons donc dans l'ère de « l'évaluation de l'évaluation », où l'évaluation devient un objet d'évaluation[485]. Étant donné que les organismes d'aide évaluent de plus en plus systématiquement leurs projets, programmes et stratégies, il devient désormais possible d'effectuer des études transversales sur l'évaluation, et ainsi d'établir une norme de qualité à partir de laquelle classer les agences d'aide.

En 2010, le CAD-OCDE a réalisé la première étude présentant des données qualitatives et quantitatives sur la manière dont les agences d'aide et les banques de développement pratiquent l'évaluation[486]. Le nombre d'évaluations menées par les agences d'aide a atteint des sommets par rapport aux décennies précédentes, en moyenne 696 par an sur la période 2005-2010 pour l'ensemble des donateurs. En moyenne, les unités centrales d'évaluation des organismes bilatéraux déclarent réaliser 8,4 évaluations par an[487]. La France et le Royaume-Uni sont bien situés par rapport à cette moyenne. Les trois organismes français chargés de l'aide au développement réalisent chacun 16 (AFD), 10 (direction générale du Trésor) et 12 (ministère des affaires étrangères) évaluations pour un budget de 6 millions de dollars, équivalent à 0,06 % du budget de l'APD, alors que ce budget atteignait seulement 1,5 millions de dollars dans les

[485] BERTHELEMY, Jean-Claude, *Sur la déclaration de Paris, Évaluer l'évaluation,* document de travail 41, FERDI, mars 2012 ; CLAPP-WINCEK, Cynthia, BLUE, Richard, *Evaluation of USAID's Recent Evaluation Experience,* USAID Working Paper n°320, juin 2011.
[486] OCDE, 2010(b), *op. cit.*
[487] Ce chiffre ne comptabilise pas les évaluations de projets, qui sont, dans la plupart des cas, décentralisées au niveau des agences de terrain.

années 1990. Nous constatons donc une forte hausse de la part budgétaire consacrée à l'évaluation, qui a quadruplé en dix ans. Au Royaume-Uni, le DFID annonce réaliser 23 évaluations par an pour un budget d'environ 9 millions de dollars, équivalent à 0,94 % du budget de l'APD, alors que ce budget était de 2 millions de dollars au début des années 1990. Ici aussi, le budget de l'évaluation a fortement augmenté, ayant plus que quadruplé. Si nous replaçons ces chiffres dans une comparaison plus globale, la France paraît allouer peu de son budget d'APD à l'évaluation, la moyenne internationale étant de 0,16 % des budgets. Elle réalise donc davantage d'évaluations que la moyenne, mais avec un budget plus restreint. Cela s'explique par le fait que la France est le deuxième plus gros donateur mondial en 2010[488] et que son taux d'évaluation est faible par rapport au nombre d'activités menées.

Cette hausse de l'activité évaluative dans les agences d'aide a donné une légitimité aux directeurs des services d'évaluation pour que cette fonction soit institutionnalisée, c'est-à-dire qu'une stratégie d'évaluation soit développée et que des principes de mise en œuvre et d'utilisation des évaluations (politique, technique, éthique) soient formalisés dans des documents politiques officiels. Ainsi, en 2009, le directeur du département d'évaluation du DFID, Nick York, a pu convaincre les dirigeants de l'organisme de l'utilité d'élaborer une politique d'évaluation[489]. Dans ce document, l'évaluation est présentée comme un instrument d'amélioration de l'efficacité de l'aide en termes de réduction de la pauvreté, ce qui rejoint les principes fondamentaux des OMD et de la Déclaration de Paris.

Nous constatons que l'approche de l'évaluation du DFID a complètement changé depuis les années 1990. Aujourd'hui, la politique d'évaluation de l'organisme se concentre sur les résultats de baisse de la pauvreté portant sur des cibles au niveau individuel[490]. Nous retrouvons ici l'approche de l'évaluation développée par les tenants de l'approche expérimentale. Cela contraste avec l'approche holistique des années 1990 qui soutenait des changements sociétaux globaux pour réduire la pauvreté. Le département d'évaluation insiste sur la notion d'indépendance des évaluations et sur la nécessité d'améliorer leur qualité. Les fonctions de redevabilité et de mesure de la performance de l'évaluation sont ainsi mises en avant. Les aspects partenariaux de l'évaluation sont également soulignés : avec les pays bénéficiaires de l'aide, avec les organismes de développement partenaires et entre les ministères britanniques chargés de la politique étrangère, dans un souci de cohérence. Le document de politique d'évaluation cherche aussi à légitimer l'importance donnée à l'évaluation en démontrant sa

[488] Statistiques du CAD-OCDE, http://www.oecd.org/fr/cad/stats/50060373.pdf, consulté le 15 novembre 2014.
[489] DIFID, 2009, *op. cit.*
[490] Excepté lorsqu'il s'agit d'évaluations d'impact.

contribution et son utilité pour l'amélioration de l'efficacité de l'aide au développement. Enfin, la réflexion sur la place de l'évaluation au sein du DFID et sur son articulation avec les autres départements de l'organisme, c'est-à-dire sur la gouvernance de l'évaluation, est davantage développée[491]. Cela montre la volonté du département d'évaluation d'ancrer et de stabiliser davantage son statut dans l'agence d'aide. La création d'une politique interne d'évaluation est le symbole d'une institutionnalisation de l'évaluation au sein de l'agence.

En France également, la gestion de l'aide au développement est de plus en plus régulée par une mesure de la performance et une évaluation régulière. L'AFD est soumise à un « contrat d'objectifs et de moyens » triennal qui fixe les allocations budgétaires de l'organisme en fonction des objectifs poursuivis. Le renouvellement, l'augmentation ou la baisse du budget est ensuite corrélé à l'atteinte de ces objectifs[492]. Selon un directeur de division à l'AFD, « *ce n'est pas en mettant une pression contractuelle sur le personnel de l'AFD que l'on va motiver les individus à bien travailler. Ce n'est pas le contrat qui motive le personnel des agences d'aide à faire un travail utile et efficace. Cette approche contractuelle et dépolitisée montre l'imprégnation de la culture anglo-saxonne dans la gestion des politiques publiques françaises* »[493].

Le document stratégique de la coopération au développement française[494] de 2011 précise que les organismes chargés de l'aide doivent rendre un rapport annuel sur leurs activités et un rapport d'ensemble sur la mise en œuvre de la politique française de coopération au développement tous les deux ans. De plus, pour tous les organismes d'aide au développement : « des indicateurs synthétiques sur les résultats prévus et obtenus seront mis au point. Un tableau d'indicateurs rendant compte des effets attendus des programmes financés aux niveaux bilatéral, communautaire et multilatéral et fournissant une appréciation sur les résultats obtenus à l'issue de leur mise en œuvre sera mis en place et permettra de communiquer de façon simple et explicite sur l'action du gouvernement »[495]. Dans cet engagement transparaît explicitement un usage de l'évaluation en termes de redevabilité. L'évaluation n'est jamais abordée comme un instrument stratégique de prospective sur les réformes à apporter au système de l'aide

[491] L'unité d'évaluation a ensuite jugé nécessaire de réviser cette politique en 2013 pour la faire mieux correspondre aux enjeux émergents de l'aide au développement, notamment le lien entre aide au développement et sécurité dans les États fragiles . Dans ce nouveau document, des approches évaluatives spécifiques à suivre auprès des États fragiles sont développées, alors que l'emphase mise dans la première version de la politique sur l'évaluation de la lutte contre la pauvreté est estompée. Nous retrouvons ainsi, à travers l'évolution de cette politique, les changements observés au niveau multilatéral dans la première partie du chapitre.
[492] Conseil d'État, *Décret n° 2009-618 du 5 juin 2009 relatif à l'Agence française de développement*, article 2, juin 2009.
[493] Entretien avec un directeur de département de l'AFD, 5 décembre 2014.
[494] Ministère des affaires étrangères, 2007, *op. cit.*
[495] Ministère des affaires étrangères, *Coopération au développement : une vision française*, document cadre, direction générale de la mondialisation, 2011.

français. Selon le CAD-OCDE, « le lien entre les résultats et les moyens alloués est encore peu probant dans la gestion de l'APD (française). Les évaluations ne sont pas utilisées comme un instrument stratégique. Une fois les cadres d'indicateurs remplis et les résultats présentés, le travail s'arrête et ne dépasse pas le niveau de la redevabilité. Les leçons tirées pour améliorer la performance de la politique française de coopération restent faibles »[496]. Ce constat, aussi vrai pour la France que pour le Royaume-Uni, confirme notre idée selon laquelle les évaluations sont peu utilisées pour répondre à des questionnements sur le futur de l'aide au développement et pour entraîner un apprentissage, mais plutôt comme des moyens de mesure de la performance des organismes d'aide et de contrôle de leur respect des bonnes pratiques de la Déclaration de Paris.

Pourtant, les responsables politiques de l'aide au développement en France déclarent avoir besoin d'évaluations expérimentales complexes pour éclairer leur prise de décision. En effet, le ministère des affaires étrangères souligne l'effort que doivent réaliser les organismes d'aide pour mettre en œuvre des évaluations expérimentales : « des évaluations d'impact seront menées selon une approche scientifique renforcée. La France soutiendra le développement de travaux permettant de rendre compte de façon objective et mesurable des effets économiques et sociaux des politiques de coopération et veillera à les utiliser pour tirer des enseignements des actions qu'elle finance »[497]. Cette citation montre que pour les responsables politiques la valeur de l'évaluation réside dans sa complexité et sa rigueur scientifique. On ne parle pas seulement d'approche « scientifique », mais d'approche « scientifique renforcée », sans que la nuance soit pour autant explicitée. L'idée implicite est que si les évaluations ne sont pas utilisées aujourd'hui dans la prise de décision politique, c'est parce qu'elles ne sont pas d'assez bonne qualité et n'apportent pas de preuves chiffrées de l'efficacité de l'aide. En ressort l'idée que les évaluations sont crédibles et utiles à la décision seulement si elles sont « objectives » et quantitatives. Cet engouement pour les méthodes expérimentales révèle aussi un enjeu « réputationnel ». En développant une expertise d'évaluation expérimentale, la France se montre à la pointe de recherche sur l'évaluation et l'efficacité de l'aide au développement.

En 2007, dans un contexte national où l'évaluation des politiques publiques gagne en importance, et sous l'influence des agences d'aide étrangères, la division de l'évaluation de l'AFD a débuté une réflexion interne, afin de mettre en place une politique d'évaluation. Cette politique devait être validée en 2008 après discussion avec les départements opérationnels et

[496] DAC-OECD, *Peer review: France*, 2013, p. 86.
[497] Ministère des affaires étrangères, 2011, *op. cit.* p. 56.

stratégiques. La conception de cette politique se fondait notamment sur les conclusions d'une étude comparative des pratiques d'évaluation d'une dizaine d'institutions de développement[498]. Toutefois, ce projet de politique d'évaluation n'a pas pu aboutir[499] et la seule résultante de l'étude a été l'adoption du principe de décentralisation. En 2012, le nouveau directeur de la division de l'évaluation, Laurent Fontaine, a relancé ce projet d'élaboration d'une politique d'évaluation, notamment sous la pression normative du CAD-OCDE, qui cette fois-ci a abouti. Le document politique est centré sur le statut, les méthodologies et l'utilité de l'évaluation au sein de l'AFD. Les responsabilités des personnels de l'organisme durant la préparation, la mise en œuvre, le suivi et la valorisation des évaluations y sont définies[500]. À travers ce projet de politique d'évaluation, la division de l'évaluation a donc clairement exprimé sa volonté d'institutionnaliser l'évaluation au sein de l'AFD, afin d'y renforcer la culture de l'évaluation. Notre observation participante à l'AFD de 2012 à 2013, nous a permis de constater que les négociations avec les dirigeants de l'agence autour de l'élaboration de ce document ont été difficiles, en raison de visions divergentes de la fonction de l'évaluation entre ces derniers et les évaluateurs. Ceux-ci voulaient un engagement des dirigeants à utiliser les évaluations dans leur prise de décision et mettant en avant la fonction d'apprentissage interne de l'évaluation, alors que les dirigeants donnaient la priorité au développement de la fonction de redevabilité de l'évaluation à moindre coût. Nous avons également interprété ces blocages comme une peur des dirigeants que l'influence de l'évaluation ne devienne trop contraignante dans la prise de décision. Un consensus a pu émerger grâce aux cadres normatifs du CAD-OCDE, sur lesquels les évaluateurs internes et le cabinet d'évaluation recruté en externe pour élaborer la politique d'évaluation se sont appuyés pour faire valoir leur approche de l'évaluation. Le CAD-OCDE est, pour illustration, cité à 24 reprises dans un document de 38 pages. Ainsi, les standards internationaux donnent un certain pouvoir aux évaluateurs pour légitimer et donner un statut à leur expertise.

Selon nous, l'émergence de ces politiques d'évaluation dans les organismes d'aide britannique et français, dans une période de temps similaire, montre que la culture de l'évaluation se diffuse dans les agences d'aide. Ces politiques ont pour objectif d'approfondir l'ancrage de cette culture et de la diffuser à tous les échelons hiérarchiques. Les discussions que nous avons pu avoir avec les évaluateurs de ces agences révèlent qu'ils attendent de ces

[498] FORESTI, Marta, *A comparative Study of Evaluation Policies and Practices in Development Agencies,* ADF-ODI, décembre 2007.
[499] Les raisons pour lesquelles cette première initiative d'élaboration d'une politique d'évaluation a échouée restent floues après notre enquête au sein de l'AFD, il est donc impossible d'en offrir une interprétation suffisamment rigoureuse.
[500] AFD, *Politique d'évaluation de l'AFD,* octobre 2013.

documents qu'ils convainquent les décideurs et les opérationnels de l'utilité de l'évaluation et qu'ils débarrassent l'évaluation de son image technocratique éloignée de la prise de décision et des enjeux du terrain[501]. C'est pourquoi les unités d'évaluation ont particulièrement tenu à ce que les politiques d'évaluation soient élaborées de manière participative avec les différents départements des agences d'aide. La phase d'élaboration des politiques a ainsi permis un dialogue interne autour de l'évaluation et une sensibilisation des différents départements opérationnels et stratégiques aux enjeux de l'évaluation. À l'AFD par exemple, des ateliers de travail ont été réalisés avec des représentants des départements stratégiques et opérationnels pour échanger sur leurs attentes en matière d'évaluation.

L'élaboration de documents politiques institutionnalisant le mandat des unités d'évaluation des agences d'aide au développement nous apparaît toutefois tardive. Alors que les agences d'aide pratiquent l'évaluation depuis les années 1960 ou 1970, il aura fallu attendre plus de trente ans pour que cette activité soit formalisée par une stratégie et des règles internes légitimées au niveau décisionnel. La formulation de politiques d'évaluation permet de sécuriser la place de l'évaluation dans les agences d'aide.

Nous allons désormais nous intéresser au travail concret des évaluateurs et étudier comment ils ont adapté leurs pratiques aux bonnes pratiques du CAD-OCDE et aux attentes des décideurs des organismes d'aide, maintenant formalisées et institutionnalisées dans des documents de stratégie politique.

- **Des méthodologies en voie de stabilisation et d'appropriation**

Dans une culture de l'évaluation où la redevabilité et la mesure de la performance sont centrales, la garantie de l'indépendance des évaluations prend de l'importance. En effet, la légitimité des évaluations dépend de leur indépendance. Cet enjeu a toujours été présent, mais s'exprime plus fortement dans les années 2000. L'AFD et le DFID ont mis en place, à la fin des années 2000, des « comités indépendants d'évaluation » ayant pour objectif de garantir l'indépendance et la qualité des évaluations. Ces comités valident les programmes d'évaluation et en contrôlent la méthodologie, les résultats et les recommandations. Les unités d'évaluation se trouvent donc supervisées par un organisme externe, composé de membres des tutelles

[501] Entretien avec une évaluatrice du DFID, 4 juillet 2014 ; conversations informelles avec des évaluateurs de l'AFD lors de notre observation participante à l'AFD en 2012 et 2013.

financières et de personnalités indépendantes qualifiées en évaluation, qui rend directement compte au conseil d'administration des agences (ou équivalent).

Cette recherche d'indépendance va de pair avec une externalisation de plus en plus grande des évaluations, notamment en réponse aux recommandations du CAD-OCDE. Si au Royaume-Uni, l'externalisation était déjà généralisée dans les années 1990, le processus a été plus tardif en France, où il s'est seulement généralise dans les années 2000.

Le DFID et deux des organismes français d'aide au développement (l'AFD et le ministère des affaires étrangères) ont également renforcé la décentralisation de la fonction de l'évaluation dans les années 2000. La décentralisation permet, tout d'abord, d'augmenter le nombre d'évaluations de projets réalisables et ainsi de faire que la quasi-totalité des projets soit évaluée, ce qui répond au principe de systématisation de l'évaluation, cher à l'idéologie du New Public Management. Toutefois, un directeur de division à l'AFD estime que « *l'externalisation, la création d'un comité des évaluations et la décentralisation des évaluations nuisent à la mise en œuvre d'évaluations substantielles dans un objectif d'apprentissage et de recherche. L'évaluation devient un pur outil gestionnaire* »[502]. En effet, au début des années 2000, l'AFD systématise l'utilisation des cadres logiques pour se conformer à la méthode de gestion de l'aide inspirée de la « théorie du changement » qui prévaut alors dans le système de l'aide au développement. L'évaluation est, dès lors, perçue comme une étape du cycle de gestion des projets, visant à retracer la chaîne de causalités ayant menée aux résultats des activités financées par l'AFD[503].

Les évaluateurs des unités centrales des organismes d'aide sont chargés de superviser les évaluations menées par le personnel des agences de terrain et de les conseiller sur le recrutement des experts, la méthodologie et la rédaction des rapports, ainsi que de valider la qualité des rapports. Cette tâche mobilise beaucoup de ressources humaines dans les unités d'évaluation et est très chronophage. Chaque évaluateur se voit assigner plusieurs agences décentralisées qu'il doit assister dans la mise en œuvre des évaluations de projets. Les documents de politique d'évaluation du DFID et de l'AFD insistent beaucoup sur les bénéfices de la décentralisation des évaluations en matière de diffusion d'une culture de l'évaluation à tous les niveaux de l'organisation et d'implication des partenaires locaux dans les évaluations. En 2011, l'AFD a produit une évaluation de son expérience de décentralisation des évaluations de projets pour

[502] Entretien avec un directeur de département de l'AFD, 5 décembre 2013.
[503] Propos d'une économiste de l'AFD recueillis lors d'un séminaire interne à l'AFD sur l'approche par les résultats que nous avons contribué à organiser en décembre 2014, voir liste des évènements observés en annexe 9.

vérifier que ce mode de fonctionnement ne contrevenait pas à la qualité des évaluations et favorisait une meilleure participation des partenaires et des experts locaux. Nous voyons, ici encore, que les donateurs réalisent de plus en plus d'évaluations de leurs pratiques d'évaluation.

Concernant le critère de participation des pays bénéficiaires aux évaluations décentralisées, il ressort de l'évaluation que 60 % de ces évaluations sont réalisées par des bureaux d'études locaux ou d'un pays voisin. La qualité des évaluations décentralisées, quant à elle, est jugée satisfaisante dans 78 % des cas. Toutefois, si l'évaluation conclut que les évaluations apportent généralement une bonne analyse de la pertinence et de la qualité de la mise en œuvre des projets, il apparaît qu'elles apportent moins d'éléments sur l'appréciation de leurs résultats et de leurs impacts[504]. Elles ne suffisent donc pas dans un objectif de redevabilité. Au DFID, les évaluateurs du Siège jugent plus favorablement la qualité des évaluations décentralisées[505].

Au DFID, à la différence de la France, les évaluations de projets sont systématiquement rendues publiques. En effet, en France, l'accord des pays bénéficiaires est nécessaire à la publication des évaluations de projets, or ces derniers s'y opposent régulièrement, notamment lorsque les rapports soulignent des dysfonctionnements de la gouvernance locale[506]. Les autres types d'évaluation (stratégiques, d'instrument, sectorielles) sont censés être rendus publics.

Cependant, même pour ces évaluations, des compromis doivent parfois être trouvés sur le format de la publication (par exemple, avec une publication qui se limite à une synthèse des conclusions, lorsque les pays bénéficiaires ou les bailleurs de fonds partenaires concernés par l'évaluation ne souhaitent pas voir le rapport divulgué publiquement dans son intégralité. Nous avons observé ce cas à l'AFD lorsque nous avons participé à l'évaluation des effets environnementaux et sociaux d'un projet de grand barrage au Laos. Ce barrage ayant fait l'objet de fortes contestations de la part d'ONG internationales de protection de l'environnement, ses effets sur la faune et la flore et les conditions sociales des populations locales étaient des enjeux particulièrement sensibles. En conséquence, les donateurs multilatéraux partenaires de l'AFD se sont opposés à la publication d'un rapport d'évaluation sur le sujet. L'AFD a décidé en conséquence, de ne pas qualifier le rapport « d'évaluation » mais de « revue interne », et n'a publié qu'un document de quatre pages résumant les « leçons apprises pour l'AFD » à partir de la revue de ce projet de barrage. Cette anecdote illustre bien les tensions politiques qui peuvent

[504] Évaluation confidentielle consultée en interne à l'AFD.
[505] Entretien avec une évaluatrice du DFID, 4 juillet 2014.
[506] Entretien avec une évaluatrice de l'AFD, 25 juin 2014.

entourer les évaluations lorsqu'elles portent sur des enjeux politiquement sensibles et la pression que peuvent exercer les donateurs les plus puissants pour garder la main sur la mise en œuvre et la publication des évaluations.

La Déclaration de Paris encourage également les donateurs à mener des évaluations partenariales avec les pays bénéficiaires de l'aide. Toutefois, cette pratique reste rare. De la même manière, les donateurs sont souvent réticents à mener des évaluations conjointes avec d'autres donateurs. Une enquête du CAD-OCDE, menée en 2010, sur les pratiques évaluatives des donateurs montre que seulement 15 % des évaluations sont menées de manière partenariale. Environ 25 % des donateurs déclarent associer les évaluateurs du Sud aux phases de décision et de planification des évaluations, alors que moins de 20 % les font participer au suivi des résultats des évaluations[507]. Selon cette même étude, les acteurs du Sud sont les plus sollicités lors des phases de collecte de données et de missions de terrain (par environ deux-tiers des donateurs) et dans une moindre mesure de la conception des évaluations (par environ 40 % des donateurs). Luc Lefebvre, directeur d'un cabinet d'évaluation luxembourgeois, déclare par exemple : « l'expertise des évaluateurs du Sud est avant tout nécessaire dans les phases de mission de terrain (entretiens, collectes de données), car ils connaissent le contexte socio-politique et peuvent faciliter le dialogue et l'accès à certaines parties prenantes locales. Ils peuvent nous permettre de percevoir des enjeux qui nous auraient échappés, car nous ne connaissons pas les codes locaux »[508]. En dehors de ces phases, les évaluateurs du Sud sont très peu intégrés aux évaluations[509].

De plus, l'inclusion aux évaluations des officiels locaux et des bénéficiaires des activités d'aide est souvent superficielle. Toutefois, des innovations intéressantes sont mises en œuvre pour améliorer l'appropriation des résultats des évaluations par les pays bénéficiaires. Ainsi, l'AFD a expérimenté les « évaluations filmées » dans le cadre d'un projet de rénovation du réseau hydraulique de Soweto. À la place du rapport d'évaluation, un film a été réalisé, présentant les réalisations et les résultats du projet, les changements qu'il a apportés dans la vie des bénéficiaires et les témoignages des différentes parties prenantes. Le format du film est censé être plus accessible aux populations bénéficiaires que celui du rapport et permettre une meilleure transmission des messages de l'évaluation par l'image. Les bénéficiaires ayant visualisé l'évaluation filmée du programme de rénovation du réseau hydraulique de Soweto

[507] *Ibid.*
[508] SFE-EVALED, *atelier de travail : associer les professionnels locaux dans les évaluations des interventions d'appui au développement – un plus ou une obligation ?* Jeudi 30 janvier 2014, voir liste des évènements observés.
[509] Nous analyserons plus en profondeur les raisons de cette difficulté qu'ont les évaluateurs du Nord et du Sud à collaborer dans le chapitre 6.

étaient enthousiastes et semblaient s'être appropriés les « leçons apprises » de l'évaluation. La parole des interviewés n'étant pas retranscrite, elle apparaissait moins instrumentalisée. Le public a apprécié la diversité des points de vues exprimés et l'absence de l'intermédiaire de « l'expert » dans la transmission de l'évaluation, même s'il a été émis que le montage du film pouvait modifier la réalité exprimée par les acteurs. En revanche, cette expérimentation a été largement critiquée par le comité des évaluations de l'AFD qui a jugé que le format du film enlevait la rigueur et l'objectivité d'un rapport écrit, que la durée du film obligeait à faire des raccourcis explicatifs et que l'on ne voyait pas assez apparaître la structure et les standards de qualité de l'évaluation[510]. Ainsi, cette expérimentation reste, aujourd'hui, unique et la direction de l'AFD est réticente à sa reproduction.

Les évaluations conjointes avec d'autres donateurs constituent, quant à elles, seulement 7 % des évaluations. Lorsqu'ils sont interrogés sur leurs expériences en matière d'évaluation conjointes, les évaluateurs expriment des difficultés à coordonner leurs calendriers, à harmoniser leurs attentes et à partager les responsabilités de chacun dans la mise en œuvre de l'évaluation. Selon un évaluateur retraité du ministère de la coopération, « *ces exercices d'évaluation conjointe sont finalement tellement lourds et longs à mener que seuls les services d'évaluation des organismes d'aide sont impliqués, et qu'on ne peut pas mobiliser les services opérationnels. Ces derniers ne peuvent donc pas s'approprier les recommandations. On est dans une stratégie « quasi-intégrée » sans capacité de tirer des enseignements* »[511]. Un membre du réseau d'évaluateurs du CAD-OCDE souligne que malgré les avantages des évaluations conjointes, il y a également de nombreuses difficultés à leur mise en œuvre : « *on ajoute de la complexité et les évaluations prennent forcément plus de temps à exécuter puisqu'il faut coordonner les évaluateurs et les organismes d'aide en permanence. Les coûts de transaction totaux sont plus élevés pour les bailleurs. L'exigence de reddition de comptes individuelle reste, car le management de chaque agence souhaite avoir des informations sur sa propre contribution* »[512]. Selon Pascal Brouillet, « *en raison de la multiplicité des donateurs impliqués, les évaluations conjointes sont très technocratiques et lourdes administrativement. La question est aussi celle du partage des responsabilités dans l'évaluation qui est toujours délicate à*

[510] Lors de notre observation participante à l'AFD, nous avons pu accéder à la documentation interne à l'AFD sur l'expérimentation de l'évaluation filmée, à des réunions de travail sur cette expérimentation et à une projection de l'évaluation filmée. À notre connaissance, cette expérimentation n'a pas été reproduite par d'autres organismes de développement.
[511] Michael Ruleta, consultant indépendant en évaluation, ancien évaluateur au ministère des affaires étrangères, lors d'un groupe de travail de la SFE en mars 2012 sur les évaluations conjointes.
[512] Hans Lundgren, président du réseau d'évaluateurs du CAD-OCDE, lors d'un groupe de travail de la SFE en mars 2012 sur les évaluations conjointes.

négocier »[513]. Nous pensons, qu'au-delà de ces difficultés, les problèmes principaux qui se posent à la mise en place d'évaluations conjointes tiennent premièrement du fait que les programmes n'ont pas été mis en place de manière conjointe par les donateurs. La plupart du temps, la mise en œuvre est fragmentée, chaque donateur se chargeant d'une thématique spécifique du programme d'aide. Une collaboration réelle dans la mise en œuvre est une condition primordiale d'une bonne évaluation conjointe. D'autre part, nous observons que les bailleurs de fonds réalisent souvent un quota d'évaluations conjointes comme l'exigent les organismes multilatéraux mais sans réelle volonté politique. Cela conduit à des stratégies de faux-semblants où les donateurs feignent de travailler ensemble.

En 2010, les évaluateurs étaient donc loin d'avoir atteint les objectifs de la Déclaration de Paris qui fixaient à 66 % le taux d'analyses conjointes ou participatives. Ces faibles résultats dans la pratique contrastent beaucoup avec les discours et les documents stratégiques des agences d'aide qui mettent très en avant l'ambition partenariale de leur approche évaluative. Nous en concluons que la conformité affichée aux standards d'évaluation du CAD-OCDE par les donateurs relève encore aujourd'hui en partie de l'affichage.

Les agences d'aide au développement sont également dans un positionnement ambigu quant à la bonne pratique des évaluations d'impact expérimentales. Cette méthodologie est toujours considérée comme le standard d'or de l'évaluation et les agences d'aide ont la volonté de développer une expertise permettant d'en faire un axe de leurs travaux d'évaluation. Des ressources humaines et des fonds importants sont consacrés au développement des évaluations d'impact, et c'est un des thèmes les plus importants d'échange d'expérience et de collaboration internationale entre les donateurs. Toutefois, les entretiens que nous avons menés nous ont permis de constater que les évaluateurs et les dirigeants des agences d'aide commencent à exprimer une certaine désillusion quant à la « révolution » qu'étaient censée apporter cette méthode dans l'analyse de l'efficacité de l'aide, rejoignant ainsi les critiques des chercheurs sur le développement évoquées ci-dessus[514].

À la fin des années 2000, en France et au Royaume-Uni, les bailleurs de fonds ont exprimé le désir d'avoir un retour sur les investissements réalisés pour le développement des évaluations d'impact expérimentales. Les évaluateurs de l'AFD pour la France et les experts de l'ODI pour le Royaume-Uni ont alors mené des études faisant le bilan des évaluations

[513] Pascal Brouillet, ancien évaluateur à l'AFD, lors d'un groupe de travail de la SFE en mars 2012 sur les évaluations conjointes.
[514] Observation participante à l'AFD et entretiens menés auprès d'évaluateurs des autres agences d'aide française et britannique en 2012 et 2013, voire liste des personnes interrogées dans ces organisations en annexe 8.

expérimentales réalisées et de l'utilisation qui avait été faite de ces évaluations[515]. Dans l'étude de l'ODI, les évaluateurs du DFID s'expriment tout d'abord sur le caractère politique de leur engouement pour la méthode expérimentale : « les évaluations expérimentales sont réalisées sous la pression des politiques et des donateurs puissants, la Banque mondiale en particulier, afin d'améliorer l'efficacité de l'aide »[516]. Les évaluateurs font donc face à une double pression : celle des décideurs politiques nationaux qui veulent disposer de preuves rigoureuses et chiffrées pour justifier leur stratégie d'aide au développement et celle des organismes multilatéraux qui ont établi l'évaluation expérimentale comme le standard d'or de l'évaluation. Selon un évaluateur de l'AFD, « *il y a aussi un enjeu économique et de pouvoir fort pour les experts des organismes multilatéraux et de structures de recherche nationales ou internationales. En effet, alors qu'en France, une évaluation « classique » coûte au maximum 80 000 euros, le prix s'élève à 600 000, voire 800 000, euros pour une évaluation expérimentale d'impact* »[517]. Les évaluations expérimentales sont donc très rémunératrices pour les évaluateurs externes qui les réalisent. Par ailleurs, les évaluations expérimentales étant le standard d'or des évaluations, il y a aussi un enjeu en termes de prestige pour les chercheurs et experts qui réalisent ces évaluations, car elles ont un retentissement fort et sont régulièrement publiées dans des revues scientifiques réputées au niveau international[518].

Le fait que les évaluations expérimentales aient comme objectif premier de prouver les bons résultats de l'aide au développement entraîne un manque de transparence dans la publication des évaluations, comme le constate l'ODI : « les résultats négatifs sont rarement publiés. Dans notre base de données, aucune évaluation ne démontre « aucun impact » ou des effets négatifs. Cela affecte le potentiel apprentissage à partir de ces évaluations et fausse la redevabilité »[519]. Nous tenons à relativiser ces propos. Dans certains cas, les impacts négatifs des programmes révélés par les évaluations expérimentales ont été abondamment discutés et diffusés, par exemple, dans le cas de l'évaluation expérimentale du programme de microcrédit de l'AFD au Maroc[520].

Nous estimons que les espoirs placés dans les évaluations expérimentales, vues comme la solution pour enfin prouver l'efficacité de l'aide, sont trop ambitieux. Comme nous l'avons

[515] NAUDET, Jean-David, DEALRUE, Jocelyne, BERNARD, Tanguy « Évaluations d'impact, un outil de redevabilité ? Les leçons tirées de l'expérience de l'AFD », *Revue d'économie du développement*, n°4, vol. 16, 2012, pp. 27-48; ODI, *Improving impact evaluations production and use,* Working Paper, mars 2009.
[516] ODI, 2009, *op. cit.*
[517] Entretien avec un évaluateur de l'AFD spécialisé dans les méthodes expérimentales, 19 novembre 2013.
[518] *Ibid.*
[519] ODI, 2009, *op. cit.*
[520] GUERIN, Isabelle, *et al, Analyse des déterminants de la demande de services financiers dans le Maroc rural*, AFD, département de la recherche et de l'évaluation, 2012.

évoqué précédemment, ce type d'évaluation ne peut que prouver l'efficacité de projets au niveau micro-économique et leur valeur ajoutée en termes d'analyse des causes de la réussite ou de l'échec des projets est faible. C'est cette surestimation qui a causé une désillusion chez les donateurs.

Dans les évaluations expérimentales réalisées par l'AFD qui ont fait l'objet d'un bilan[521], la méthodologie stricte de l'expérimentation aléatoire (la constitution randomisée de groupes test et témoins de bénéficiaires) n'a pas pu être mise en place pour des raisons opérationnelles (faible taux de participation, modification du projet en cours de route, abandons de la participation à l'expérimentation, etc.). Par exemple, les évaluateurs ayant réalisé l'évaluation d'un projet d'assurance maladie au Cambodge racontent : « un certain nombre de difficultés sont apparues. Les taux d'abandon ont été importants au bout de la période couverte par les bons de réduction (pour accéder à l'assurance maladie), ce qui a entraîné sa prolongation. L'étude voulait en plus apprécier la manière dont l'assurance contribuerait à réduire l'endettement lors d'un choc sanitaire grave – un événement trop rare pour qu'il soit statistiquement décelable. Enfin, les données collectées sur la consommation de services de santé étaient trop réduites pour que l'on puisse en tirer des mesures significatives d'une évolution des comportements et l'amélioration des performances sanitaires identifiée a été très faible »[522]. Ces difficultés contextuelles, que l'on pourrait retrouver dans d'autres projets d'aide, ont empêché le bon fonctionnement de la méthodologie expérimentale.

Au sujet de la méthodologie, l'ODI recommande aux évaluateurs « de ne pas être trop orthodoxes sur la rigueur des procédures et de prendre en compte les difficultés inhérentes au contexte de l'évaluation, où les données et les capacités statistiques sont souvent relativement faibles, étant donné le manque de moyens des administrations dans les pays bénéficiaires de l'aide »[523]. Or, étant donné que toute la valeur ajoutée des évaluations expérimentales provient de la rigueur de leur méthodologie, apporter de la souplesse à cette rigueur apparaît problématique.

Les évaluateurs de l'AFD et de l'ODI arrivent à la même conclusion : les évaluations expérimentales ne sont appropriées que pour un ensemble restreint de projets, qui regroupent les caractéristiques suivantes :

---

[521] LEVINE, David, POLIMENI, Rachel, RAMAGE, Ian, *Insuring Health or Insuring Wealth? An Experimental evaluation of Health Insurance in Cambodia*, AFD, département de la recherche et de l'évaluation, 2012 ; GUERIN, Isabelle, *et al*, 2012, *op. cit.*
[522] *Ibid*.
[523] ODI, *Improving impact evaluation production and use*, Working Paper, mars 2009, p.12.

- des réalisations circonscrites et homogènes ;
- une mise en œuvre éprouvée et une chaîne causale courte et étanche aux facteurs extérieurs au projet ;
- une participation large et durable et des comportements stables ;
- des effets rapides et à la dynamique simple.

Pour décrire ces projets, ils utilisent l'analogie d'un tunnel « qui a un début et une fin délimités, où tout ce qui entre peut facilement être défini comme un intrant et tout ce qui sort comme un produit, où il est pratiquement impossible d'abandonner en cours de route, et où le parcours est à la fois court et prévisible de sorte qu'il est préservé de toute influence extérieure. Dire que les évaluations d'impact sont le standard d'excellence de l'évaluation revient à dire que le marteau est le standard d'excellence des outils, cela n'est vrai que si l'on souhaite enfoncer des clous »[524].

La volonté qu'ont la plupart des évaluateurs et des décideurs politiques de voir ces méthodes appliquées au plus grand nombre de projets de développement n'a donc pas de sens. En effet, les projets des agences d'aide, ne répondent que très rarement aux critères nécessaires à la mise en œuvre d'évaluations expérimentales. Les évaluateurs se voient alors contraints de modifier la méthodologie et amoindrissent sa rigueur scientifique. Les évaluations perdent alors leur valeur ajoutée en termes de redevabilité. De plus, le type d'interventions qui peut être évalué de manière expérimentale ne correspond plus aux bonnes pratiques internationales de gestion de l'aide. Il s'agit de projets dirigés par les donateurs (puisqu'ils nécessitent la création de groupe de contrôle, l'adaptation des projets aux contraintes de l'évaluation expérimentale, etc.) et non par les bénéficiaires. Ce sont des projets de petite taille ayant des chaînes causales simples et non des projets multisectoriels avec des effets divers et complexes sur le développement. Ils se fondent souvent sur des indicateurs purement économétriques, ce qui s'oppose à la reconnaissance de la nature multidimensionnelle de la pauvreté et du développement. Enfin, leur structure est rigide et ne permet pas une adaptation aux évolutions sociétales en cours de mise en œuvre.

En outre, nous estimons que les méthodes expérimentales d'évaluation affaiblissent le potentiel d'apprentissage mutuel entre donateurs et bénéficiaires que peuvent avoir les évaluations. Selon un évaluateur à l'AFD spécialisé sur les méthodes expérimentales : « *il y a aujourd'hui une vraie négation des savoirs pratiques par l'évaluation expérimentale. Il y a une*

[524] ODI, 2009, *op. cit.* p. 6. *;* NAUDET, Jean-David, DELARUE, Jocelyne, 2012, *art. cit.* p.46.

impression forte de négation de ce que font les experts et professionnels locaux. Or, ces gens-là ont une vraie culture dans leur domaine et accumulent du savoir et de l'expertise technique depuis des décennies, et on vient leur dire que « ça, ce n'est pas de la science », la domination de l'expert est alors très forte »[525].

Nous pouvons donc conclure que la méthode expérimentale, loin de révolutionner les politiques sociales, doit être croisée à d'autres méthodes d'évaluation pour apporter une vision complète des processus de développement. En effet, chaque méthodologie incarne une vision particulière du développement, reflétée par les débats académiques constants pour en imposer l'une par rapport à l'autre.

- **Conclusion de chapitre**

Aussi bien les OMD que la Déclaration de Paris ont donné une place centrale à l'évaluation comme outil de mesure de la performance de l'aide et d'allocation « rationnelle » des fonds. L'obtention de résultats est aujourd'hui ce qui légitime les politiques et les pratiques d'aide au développement. C'est pourquoi les décideurs politiques ont des attentes de plus en plus fortes envers les évaluateurs pour qu'ils apportent des preuves scientifiques des résultats des projets et des programmes qu'ils mettent en œuvre. En résulte un développement des méthodes expérimentales d'évaluation et une utilisation des évaluations tournée vers la redevabilité, tendance qui n'est pas favorable à l'apprentissage, objectif pourtant fondamental de l'évaluation car il permet une accumulation de connaissances sur les processus de développement. L'évaluation devient un outil de « gestion de l'aide » et perd peu à peu ses fonctions d'instrument de « recherche » et de « prospective » sur l'aide au développement.

Un point important de notre argumentation est que cette volonté de quantifier les résultats de l'aide au développement, de les suivre et de les évaluer selon des méthodologies complexes et « scientifiques », traduit une volonté des bailleurs de fonds de donner à l'aide au développement l'image d'une « technique » plutôt que d'une « politique ». Dépolitiser, en façade, l'aide au développement permet aux décideurs de se protéger de la critique qui pourrait remettre en cause les choix d'allocation de l'aide ou de stratégies de développement déterminés pour des raisons politiques qu'elles soient historiques, diplomatiques ou économiques. Le fait

[525] Entretien avec un évaluateur de l'AFD spécialisé dans les méthodes expérimentales, 19 novembre 2013.

que ces choix résultent d'un suivi-évaluation rigoureux et scientifique les rend effectivement plus difficilement attaquables.

Ce chapitre met également en évidence une systématisation, une externalisation et une décentralisation de la fonction d'évaluation, répondant aux principes du New Public Management. Ce contrôle continu et indépendant de l'aide au développement, à tous les échelons de sa mise en œuvre, relève également d'une volonté des acteurs de l'aide, de gérer l'incertitude des processus de développement, ainsi que de se protéger des risques d'échecs qui accompagnent cette incertitude.

Ici s'achève la première partie de notre recherche. Cette approche socio-historique nous a permis de contextualiser, de politiser et de donner du sens à la pratique de l'évaluation telle qu'on l'observe aujourd'hui dans les organismes d'aide au développement. Étudier les évolutions de l'évaluation sur une cinquantaine d'années et les relier aux changements successifs des politiques d'aide au développement, offre une histoire politique de l'évaluation qui diffère du découpage temporel technique qu'ont pu en faire certains chercheurs.

Plusieurs tendances sont également observables. Dans les années 1960 et 1970, les évaluations sont perçues comme porteuses d'une vérité sur l'aide au développement et sont réalisées de manière ponctuelle pour apporter un apprentissage sur des projets expérimentaux. La croyance forte en l'efficacité de l'aide fait que l'évaluation est rarement menée. Dans les années 1980, cette vision d'une évaluation objective et scientifique perdure mais elle est utilisée pour prouver l'efficacité de la politique du Consensus de Washington dans un contexte où cette efficacité est de plus en plus remise en cause. Dans les années 1990, nous notons, pendant un temps, un engouement pour les approches qualitatives de l'évaluation qui « donnent la parole aux bénéficiaires », que nous interprétons comme une réaction à l'échec du Consensus de Washington. Les acteurs de l'aide perçoivent qu'ils doivent apprendre des pays du Sud pour être efficaces. Toutefois, la diffusion de l'idéologie du New Public Management dans les organismes d'aide coupe court à cet engouement et donne place à une approche gestionnaire de l'évaluation fondée sur la mesure quantitative des résultats, tendance qui s'affirmera jusqu'en 2010 pendant toute l'ère de « l'efficacité de l'aide ». À partir de 2010 et de la fin de l'agenda de l'efficacité de l'aide, nous percevons un retour des approches qualitatives, ainsi que de la fonction d'apprentissage et d'égalitarisation des rapports entre donateurs et bénéficiaires des évaluations. Nous voyons ainsi que l'évaluation évolue de manière cyclique oscillant entre des méthodes qualitatives et quantitatives et entre des fonctions d'apprentissage, de contrôle et de redevabilité.

PARTIE II : La diffusion d'une culture de l'évaluation : acteurs, leviers et réseaux.

Nous avons présenté, dans la première partie de la recherche, une histoire politique de la diffusion d'une culture de l'évaluation dans le secteur de l'aide au développement. Nous avons montré les étapes d'élaboration et d'institutionnalisation de « bonnes pratiques » de l'évaluation dans les instances multilatérales du CAD-OCDE, de la Banque mondiale et des Nations unies. Bonnes pratiques qui ont finalement été adoptées par tous les bailleurs de fonds occidentaux. Nous avons analysé comment les évolutions de la politique internationale de l'aide au développement ont influencé les « bonnes pratiques » de l'évaluation, en termes de contenu et de méthodologie. Nous avons également, au travers des cas britannique et français, démontré le poids des cultures nationales des politiques publiques sur la manière dont les agences d'aide se sont appropriées et ont appliqué la norme internationale de l'évaluation.

Après avoir analysé le processus de long terme de construction d'une culture de l'évaluation dans le secteur de l'aide au développement, nous allons étudier sa diffusion dans différents espaces du système de l'aide. Cela nous permettra d'avoir une meilleure compréhension du rôle et de l'insertion des évaluateurs dans ce système. En effet, s'il existe des études sur les techniques d'évaluation de l'aide au développement, elles sont relativement désincarnées et n'étudient pas les acteurs qui portent et mettent en œuvre ces techniques. Nous proposons donc de « réincarner » l'évaluation et ainsi de montrer qu'elle est un objet politique composé de croyances, de standards, d'acteurs ayant des intérêts et des ambitions dans le système de l'aide.

Nous allons interroger le statut qu'ont acquis les évaluateurs dans le régime de l'aide[526]. Qui sont les évaluateurs ? Comment s'inscrivent-ils dans le régime dans lequel ils agissent ? Comment sont-ils perçus par les autres acteurs et comment interagissent-ils entre eux et avec les autres parties prenantes ? S'intéresser à la culture de l'évaluation demande d'étudier les interactions qui se nouent autour de l'évaluation. Au vu de l'importance gagnée par l'évaluation dans les pratiques d'aide, on pourrait penser qu'ils ont une légitimité grandissante et un pouvoir d'influence considérable sur les autres acteurs de l'aide au développement. Au contraire, nous

[526] Il ne s'agit pas ici d'opposer les processus "institutionnel" de normalisation des processus "interactionnels", mais de souligner le poids important des acteurs dans la diffusion d'une culture de l'évaluation dans les différents « espaces » de l'aide au développement. Poids qui est souvent sous-estimé par la littérature sur l'évaluation de l'aide au développement. Nous observerons d'ailleurs comment les dynamiques organisationnelles et celles produites par les acteurs se conjuguent pour entraîner des changements de pratiques et de politiques.

pensons que malgré cette importance donnée à l'évaluation, les évaluateurs ont peu de pouvoir dans leurs organisations. Il y aurait donc un décalage entre le statut donné à la fonction d'évaluation et le statut donné aux évaluateurs, dont il apparaît intéressant d'explorer les effets. En raison de ce décalage, les évaluateurs sont à la recherche d'une plus grande légitimité. Ce processus de légitimation passe, notamment, par la professionnalisation de l'évaluation et par la création d'une communauté internationale d'évaluateurs au Nord et au Sud. Notre étude portera sur trois espaces du système de l'aide au développement dans lesquels les évaluateurs agissent et diffusent leur expertise : les organismes d'aide, les pays bénéficiaires et les sociétés professionnelles d'évaluateurs.

Dans le processus de transfert d'une culture de l'évaluation à travers ces différents espaces, plusieurs types d'acteurs sont mobilisés. Reprenant le modèle d'Harold Wolman[527], nous identifions premièrement « les inspirateurs », c'est-à-dire les acteurs dont les pratiques sont prises pour modèle en vue d'une transposition, ainsi que ceux qui conceptualisent ces pratiques pour qu'elles deviennent « universelles ». Deuxièmement, nous identifions « les opérateurs de transfert », soit les acteurs qui prennent une part active à la transplantation de ces idées ou solutions politiques. Ils se caractérisent souvent par leur position à l'interface d'univers sociaux différents. Ensuite, nous avons « les facilitateurs », c'est-à-dire les acteurs qui contribuent indirectement au succès du transfert en créant des conditions favorables à sa réception ou en soutenant les opérateurs. Enfin, les « receveurs », qui mettent en œuvre ou utilisent la solution transférée[528]. Toute la complexité du transfert réside dans le fait que les acteurs peuvent tenir simultanément plusieurs rôles et que leurs discours et stratégies peuvent varier selon la position qu'ils adoptent. Nous allons démontrer que c'est la qualité des interactions entre ces différents niveaux d'acteurs qui détermine la réussite ou l'échec de l'importation de la culture de l'évaluation dans une organisation, que celle-ci soit nationale, internationale, gouvernementale ou de la société civile.

Dans un premier chapitre, nous analyserons l'influence que l'évaluateur a dans les organismes d'aide au développement. Nous explorerons notamment les relations qu'il entretient avec les autres acteurs de ces organismes. Nous étudierons quelles sont les pratiques, les compétences et les aspirations des évaluateurs, puis comment ces derniers perçoivent leur rôle dans ces structures. Nous analyserons également la manière dont les évaluateurs sont perçus

[527] WOLMAN, Harold, "Policy transfer among local governments: an information-theory approach", *Governance*, vol. 15, n°4, 2002, pp. 379-396.
[528] Ce modèle sera utilisé de manière plus concrète et spécifique à la politique de l'évaluation dans les chapitres successifs de la partie II.

par les autres acteurs des agences d'aide (dirigeants et opérateurs) et comment les structures et la culture organisationnelle dans lesquelles ils s'inscrivent jouent sur leur degré d'influence. Enfin, nous montrerons comment les évaluations sont utilisées par les différents acteurs. Cette étude sera basée sur les cas des organismes français d'aide au développement[529], du DFID et de la Banque mondiale[530].

Dans un deuxième chapitre, nous étudierons comment les évaluateurs des agences d'aide tentent de transférer leur expertise à des acteurs du Sud pour qu'ils deviennent des évaluateurs dans leurs pays. Nous étudierons les schémas de pensée qui fondent l'action des évaluateurs en termes de renforcement des compétences en évaluation au Sud et les instruments qu'ils mobilisent à cette fin. Nous analyserons également à la perspective des pays bénéficiaires sur ce transfert de compétences. Nous étudierons comment ces pays réagissent face à cette « solution politique » proposée par les donateurs pour améliorer leur gouvernance et dans quelle mesure ils se l'approprient. Enfin, nous remettrons en cause la logique d'importation de la culture de l'évaluation actuellement poursuivie par les donateurs occidentaux.

Dans un troisième chapitre, nous étudierons le processus de construction d'une « communauté internationale d'évaluateurs », à travers la création de sociétés internationales, régionales et nationales d'évaluation. Nous analyserons comment ces réseaux d'acteurs sont apparus et se sont multipliés et quel est leur rôle dans la diffusion d'une culture internationale de l'évaluation. Nous nous intéresserons aux instruments que ces réseaux d'évaluateurs mettent en place pour renforcer et diffuser une approche de l'évaluation normalisée dans les pays du Nord et du Sud. Nous montrerons que ces sociétés d'évaluateurs se posent en alternative à « l'importation Nord-Sud » des « bonnes pratiques » de l'évaluation *via* l'aide au développement, pour diffuser de manière plus démocratique et efficace une culture internationale de l'évaluation. Nous porterons un regard critique sur le fonctionnement de ces réseaux et le pouvoir qu'ils ont réellement sur les politiques des gouvernements du Sud et des organismes de développement.

[529] Observation participante de 2012 à 2013 au sein de l'AFD qui nous a permis d'observer les évaluateurs de cet organisme ainsi que ceux du ministère des affaires étrangères et de la direction générale du Trésor avec qui ils collaborent.
[530] Des entretiens ont été menés en 2012 et 2013 avec des évaluateurs du DFID et de la Banque mondiale (voir liste d'entretiens en annexe 8). Nos constats se fondent également sur la littérature grise produite par ces organismes.

Chapitre 4 : Les évaluateurs, des acteurs peu reconnus

Dans ce chapitre, nous allons porter le regard sur les acteurs qui incarnent la culture actuelle de l'évaluation dans les organismes d'aide : les évaluateurs. Comment agissent-ils pour faire reconnaître leur expertise et diffuser une culture de l'évaluation dans les organismes d'aide ? Comment leur expertise est-elle perçue par les autres acteurs ? Disposent-ils d'une sphère d'influence suffisante pour inciter leurs collègues à intégrer l'évaluation dans leur travail ?

Nous allons développer une analyse des interactions qui se structurent autour de l'évaluation dans les agences d'aide au développement en nous appuyant notamment sur les théories de Michel Crozier et Erhard Friedberg. Nous considérons l'évaluateur comme un « expert » qui détient un pouvoir sur l'organisation : « du fait de sa capacité personnelle à contrôler une certaine source d'incertitude affectant le fonctionnement de l'organisation »[531]. En effet, l'évaluateur détient la capacité d'expertiser l'efficacité de l'organisation par rapport aux objectifs fixés en amont des activités qu'elle met en œuvre. L'évaluateur en détenant le savoir permettant de déterminer cette efficacité devrait donc avoir un rôle et un pouvoir centraux dans le fonctionnement des agences d'aide.

Paradoxalement, les entretiens que nous avons menés et l'observation participante que nous avons réalisée à l'AFD montrent que l'évaluation et l'efficacité de l'aide sont deux champs vus comme distincts dans les agences d'aide. Un consultant en évaluation britannique déclare que « *l'efficacité de l'aide et l'évaluation sont deux activités séparées au DFID. La première porte sur la mesure des résultats, alors que la seconde vise à tirer des leçons des programmes passés. Les deux départements communiquent peu et développent des expertises distinctes* »[532]. Ce constat est vérifiable à l'AFD, où il y a peu de dialogue entre la division de l'évaluation et les divisions chargées de la production d'indicateurs de résultats. Selon une économiste de l'AFD, « *l'évaluation et la production d'indicateurs ont peu à voir l'une avec l'autre. Nous nous concentrons sur la redevabilité à court terme, alors que l'évaluation se consacre aux résultats de long terme et à l'apprentissage* »[533]. Toutefois, l'évaluation prenant de plus en plus une fonction de redevabilité au sein des agences d'aide, la réalité de cette distinction apparaît questionnable. Une confusion des rôles apparaît alors. Selon une évaluatrice de la Banque

[531] CROZIER, Michel, FRIEDBERG, Erhard, 1977, *op.cit*, p. 84-86.
[532] Entretien avec un évaluateur indépendant travaillant auprès organismes d'aide au développement, 18 février 2014.
[533] Entretien avec un économiste de l'AFD, 23 juin 2014.

mondiale, « *ces deux agendas de l'aide au développement ont toujours été traités de manière séparée, alors qu'ils touchent tous deux aux résultats des programmes d'aide. Toutefois, une synergie se crée depuis peu, au CAD et à la Banque mondiale par exemple, qui ont réuni les deux agendas dans des mêmes structures* »[534].

Cette disjonction entre l'expertise sur l'évaluation et celle sur l'efficacité de l'aide montre que le savoir des évaluateurs est perçu comme peu opérationnel (servant un apprentissage de long terme) et moins pointus par rapport à l'expertise des économistes ou statisticiens travaillant sur l'efficacité de l'aide (servant à produire régulièrement des données quantitatives sur les résultats de l'aide dans un objectif de redevabilité), et ce malgré le caractère de plus en plus « gestionnaire » que prend l'évaluation. Ainsi nous avons noté, lors de notre observation participante à l'AFD, que l'évaluation qui est, dans les faits, de moins en moins un instrument d'apprentissage, est toujours perçue comme tel par le personnel opérationnel et décisionnaire des agences d'aide, ce qui entraîne une certaine confusion. Entre chercheurs et gestionnaires, les évaluateurs peinent à trouver leur place dans les organismes d'aide.

De plus, le pouvoir de l'expert qu'est l'évaluateur est concurrencé par le pouvoir politique ou « de l'expérience » que possèdent les décideurs et les opérationnels des agences d'aide. Michel Crozier explique que dans les organisations, au « pouvoir de l'expert » se confronte le « pouvoir fonctionnel », c'est-à-dire « le pouvoir dont certains individus disposent du fait de leur fonction dans l'organisation pour contrôler le pouvoir de l'expert et, à la limite, y suppléer »[535]. Ces individus vont tenter d'instrumentaliser le savoir de l'évaluateur pour l'orienter vers le renforcement de leur propre pouvoir au sein de l'organisation. Ainsi, la capacité de l'évaluateur de diffuser et de faire valoir son expertise dépend, en partie, de son positionnement dans la structure organisationnelle. Ce positionnement reflète le degré d'importance donné par l'organisation aux évaluateurs et le rôle qui leur est assigné

Par ailleurs, les évaluateurs ont eux-mêmes une vision spécifique de leur fonction. Cette vision peut différer de celle entretenue par les autres catégories d'acteurs. Dans l'ensemble structurel, normatif et cognitif que sont les organismes d'aide, les évaluateurs vont développer des comportements et des stratégies qui reflètent cette vision. Nous rejoignons la thèse de Michel Crozier selon laquelle ce ne sont pas les structures et les objectifs des organisations qui permettent d'expliquer les comportements des acteurs en leur sein, puisque ces derniers

[534] Entretien avec une évaluatrice de la Banque mondiale, 24 octobre 2013.
[535] CROZIER, Michel, *Le système bureaucratique*, Éditions du Seuil, 1963, 369p, pp. 202.

disposent de marges de liberté pour poursuivre leurs activités[536]. C'est pourtant le biais qu'ont suivi les membres du CAD-OCDE et des agences bilatérales quand ils ont pensé qu'il suffirait de placer les départements d'évaluation à une position hiérarchique élevée et indépendante des départements opérationnels, ainsi que de mettre en place des « bonnes pratiques » d'évaluation pour que les développeurs prennent en compte l'évaluation dans leurs pratiques.

A. Une négation des dynamiques interactionnelles

Plutôt que de suivre cette approche institutionnelle, nous préférons adopter une approche de sociologie des organisations qui semble plus pertinente pour analyser l'état de la culture de l'évaluation dans les agences d'aide, en partant de l'analyse des perceptions et des logiques des acteurs. Nous allons montrer comment la négation des aspects cognitifs et sociologiques de la diffusion des bonnes pratiques de l'évaluation par les « inspirateurs » et les « opérateurs » du transfert de ces bonnes pratiques dans les organismes d'aide, a contribué à l'échec partiel de ce transfert.

▪ Une harmonisation des structures d'évaluation dans les agences d'aide

Pour renforcer leur pratique de l'évaluation, les agences d'aide ont pendant longtemps mis en place des réformes qui ne touchaient qu'aux structures d'évaluation (positionnement institutionnel, mise en place de système de suivi des recommandation, instauration de comités de contrôle qualitatifs, etc.), en laissant de côté les questions de l'utilisation des évaluations, de la sensibilisation du personnel aux bénéfices de l'évaluation, du dialogue entre les évaluateurs et les autres agents des organismes d'aide, ou encore de l'appropriation par les évaluateurs des « bonnes pratiques internationales » d'évaluation[537].

Cependant, les agences d'aide commencent depuis peu à s'intéresser au sens des évaluations, à leur appropriation et à leur utilisation en interne, c'est-à-dire aux aspects « immatériels » et « interactionnels » de la diffusion d'une culture de l'évaluation[538]. Ainsi, la

[536] CROZIER, Michel, 1963, *op cit.* pp. 175-176.
[537] Nous fondons cette observation sur l'étude des processus de réforme de la fonction d'évaluation des agences d'aide bilatérales européennes à partir des revues par les pairs du CAD-OCDE de 1980 à 2012 et d'entretiens auprès d'évaluateurs du DFID et de l'AFD pour une analyse plus approfondie sur ces deux structures.
[538] BORMANN, Axel, STOKHMANN, Reinhard, *Evaluation in German Development Cooperation, A System Review*, evaluation report 044, Federal Ministry for Economic Development Cooperation, 2010.

culture de l'évaluation s'est diffusée dans les agences d'aide d'une manière que nous qualifions de très empirique : la pratique de l'évaluation s'est systématisée, puis des structures et des normes ont été mises en place et dans un dernier temps une réflexion a été menée sur les fondements et les objectifs de l'évaluation. Nous expliquons cet enchaînement par le fait que les agences d'aide se sont d'abord attachées à afficher au plus vite leur respect des standards d'évaluation pour gagner la reconnaissance de leurs pairs et apparaître comme des organisations de confiance dans le système de l'aide au développement. Ce n'est que dans un second temps, lorsque ces bonnes pratiques se sont institutionnalisées, que les acteurs des organismes d'aide ont pris le temps de réfléchir à leur sens, au-delà de la conformité à un régime normatif international.

Ainsi, la première réforme mise en place par les organismes d'aide, dans les années 1980 et 1990 a souvent été d'isoler les évaluateurs des départements opérationnels et stratégiques pour garantir leur indépendance. En tenant des discours mettant en avant l'importance de l'indépendance des évaluations, les agences d'aide ont renforcé leur fonction de redevabilité. Alors que les évaluateurs estiment pour la plupart qu'il est possible d'allier apprentissage et redevabilité dans les évaluations, les opérationnels et les responsables des organismes d'aide opposent ces deux fonctions[539]. Par exemple, un directeur de division à l'AFD, déclare : « *les évaluations ne peuvent servir à la fois un objectif d'apprentissage et un objectif de redevabilité. Si l'évaluateur arrive dans une logique de contrôle des résultats, les agents évalués vont être méfiants et vont chercher à cacher les faiblesses des programmes dont ils ont la charge. Cette réaction naturelle va à l'encontre de l'apprentissage qui demande de réfléchir sur ces faiblesses* »[540].

Nous ne considérons pas que les fonctions de redevabilité et d'apprentissage de l'évaluation s'opposent fondamentalement. Montrer qu'un apprentissage a lieu, à partir d'évaluations expliquant les causes de la réussite ou de l'échec de programmes d'aide, pourrait être une forme de redevabilité. Toutefois, la redevabilité est aujourd'hui comprise par les dirigeants des organismes d'aide comme l'apport de preuves quantifiées de résultats des programmes d'aide. Cette approche de la redevabilité rend, en effet, difficile l'intégration d'une dimension d'apprentissage dans les évaluations qui nécessiterait une analyse qualitative du processus de mise en œuvre et des résultats des programmes.

[539] Propos recueillis lors de discussions informelles durant l'observation participante à l'AFD en 2012 et 2013.
[540] Entretien avec un directeur de département de l'AFD, 5 décembre 2013.

Selon nous, aborder l'indépendance du département d'évaluation seulement par la question de son positionnement dans l'organisation reflète une approche rationaliste de l'évaluation. Un chercheur danois analyse en des termes proches les réformes de la DANIDA en matière d'évaluation : « la DANIDA s'inspire principalement de l'image de l'organisation rationnelle dans sa présentation et sa conceptualisation de la fonction d'évaluation. Il est considéré qu'une vérité objective peut être trouvée à partir d'une évaluation indépendante et impartiale »[541]. Nous voyons ici comment la structuration de l'évaluation dans l'organisation reflète une approche politique particulière de l'évaluation. Notre recherche nous a amené à penser que le constat formulé dans cette citation à propos de la DANIDA est applicable à la plupart des agences d'aide. Cette similarité des réformes en matière d'évaluation en donne un aspect très mécaniste et désincarné.

Les organigrammes des agences d'aide montrent que la tendance générale est à l'isolement des départements d'évaluation dans l'organisation, afin que les évaluateurs ne puissent pas être accusés d'entrer en collusion avec les départements opérationnels et stratégiques[542]. Les départements d'évaluation rendent directement compte, soit à un conseil d'administration, soit aux tutelles financières de l'organisation. À cela s'ajoute la mise en place de comités externes des évaluations, chargés de garantir la qualité et l'indépendance des évaluations et même d'en mener comme c'est le cas au Royaume-Uni[543]. Notre ambition dans ce chapitre est d'étudier le ressenti des évaluateurs face à ce contrôle accru[544] de leur travail et les rapports de force qui naissent de ce ressenti.

Notre observation participante à l'AFD nous a permis de constater des oppositions de points de vue importantes entre le personnel de la division de l'évaluation et le comité de contrôle externe. Les évaluateurs expriment des difficultés à se situer par rapport à ce comité. Leur impression est « *qu'il n'est pas une enceinte de dialogue mais d'édiction d'une stratégie de l'évaluation préconçue par les membres du comité, parfois trop orthodoxes quant aux méthodologies à appliquer* »[545]. En effet, jusqu'alors, la supervision externe exercée par les tutelles financières relevait surtout du contrôle financier. Désormais, le contrôle exercé par les

[541] SCHAUMBURG-MULLER, Henrik, "Use of aid evaluation from an organizational Perspective", *Evaluation,* vol. 11, n°2, 2005, pp. 207-222.
[542] Étude réalisée à partir des organigrammes des agences d'aide au développement européenne (françaises, britannique, allemandes, suédoise, danoise), américaine et japonaise.
[543] En 2008, *l 'Independent Advisory Committee for Development Impact* (IADCI) a été créé. Il est chargé de superviser le travail de l'unité d'évaluation du DFID et de mener des évaluations sur la stratégie de d'aide au développement du Royaume-Uni.
[544] Pour un historique de l'accroissement du contrôle sur la fonction des évaluateurs internes aux agences d'aide, se référer à la première partie de la recherche.
[545] Propos recueillis lors de discussions informelles avec des évaluateurs de l'AFD lors de l'observation participante réalisée en 2012-2013.

comités des évaluations, structures de plus en plus répandues dans les organismes d'aide, prend un caractère qualitatif. Ces comités valident les programmes d'évaluation déterminés par les unités d'évaluation et critiquent la qualité des travaux réalisés. Face à ce contrôle, les évaluateurs, aussi bien de l'AFD que du DFID, déclarent ressentir « *une perte de liberté dans leur travail » et « une remise en cause de leur expertise* »[546]. Si les évaluations gagnent en indépendance, les évaluateurs se voient, quant à eux, destitués d'une grande partie de leur autonomie[547].

Nous voyons que les agences d'aide ne considèrent l'indépendance des évaluateurs que d'un point de vue organisationnel. Ingie Hovland identifie pourtant, dans une étude sur l'apprentissage organisationnel, deux autres types d'indépendance : l'indépendance comportementale et l'indépendance par rapport aux influences externes[548]. Ces deux derniers aspects semblent être occultés par les organismes d'aide.

L'indépendance comportementale renverrait à la liberté laissée aux évaluateurs de critiquer les pratiques de l'agence et de faire des recommandations pour leur amélioration. L'indépendance organisationnelle augmente-t-elle l'indépendance comportementale des évaluateurs ? Il semble que non. En effet, à la Banque mondiale, où les évaluations sont encore en grande partie internalisées, les évaluateurs ont une indépendance comportementale forte[549]. Les évaluations sont reconnues comme étant critiques par rapport au travail de la Banque, que ce soit au niveau des projets, des pays, des secteurs, ou de manière globale. Les témoignages que nous avons recueillis[550] relatent que, tout en reconnaissant les réussites de la Banque, les évaluations abordent aussi d'éventuels manquements, tirent des leçons de ces expériences et formulent des recommandations pour améliorer les interventions de l'organisation. Mais, ce sont ici les dires d'évaluateurs que nous retranscrivons, qui peuvent donc manquer d'objectivité quant au jugement de leur propre indépendance. Toutefois, la littérature grise et les articles produits sur l'indépendance des évaluateurs à la Banque mondiale tendent à confirmer leurs récits[551].

---

[546] Propos recueillis auprès d'évaluateurs de l'AFD lors de l'observation participante réalisée en 2012-2013.
[547] Entretien avec une évaluatrice du DFID, 12 février 2014.
[548] HOVLAND, Ingie, *Knowledge Management and organizational learning: an international development perspective*, ODI, août 2003.
[549] Entretien avec une évaluatrice retraitée de la Banque mondiale, 20 novembre 2013 ; Entretien avec une évaluatrice de la Banque mondiale, 24 octobre 2013 ; Entretien avec un consultant britannique en évaluation, 18 février 2014.
[550] Entretien avec une évaluatrice retraitée de la Banque mondiale, 20 novembre 2013 ; Entretien avec une évaluatrice de la Banque mondiale, 24 octobre 2013.
[551] World Bank, *Evaluation and Poverty Reduction, proceedings of a World Bank Conference*, OED, 1999; GRASSO, Patrick. G, "What Makes an Evaluation Useful? Reflections from Experience in Large Organizations", *American Journal of Evaluation*, vol. 24, 2003, n°4, pp. 507-514.

Au contraire, dans les agences bilatérales, le choix a été unanimement fait d'externaliser la fonction d'évaluation, solution promue par le CAD-OCDE, pour garantir l'indépendance comportementale des évaluateurs. Le changement a été très important pour les évaluateurs, dans la mesure où leur rôle n'est plus réellement de réaliser les évaluations, mais plutôt d'en superviser la mise en œuvre par des consultants extérieurs. La tâche des évaluateurs se réduit donc désormais à lancer des appels d'offre, sélectionner les cabinets d'évaluation ou les experts, effectuer la coordination entre les parties prenantes des programmes évalués et les évaluateurs, ainsi qu'à superviser la mise en œuvre des évaluations et valider leurs recommandations.

Les évaluateurs externes sont censés être moins soumis que les évaluateurs internes à la pression des acteurs de l'organisation dont les activités sont évaluées. Mais l'externalité à l'organisation est-elle une garantie suffisante pour assurer que les évaluateurs aient un comportement indépendant ? Les chercheurs qui ont travaillé sur l'évaluation de l'aide au développement sont plutôt sceptiques à ce sujet. Katharina Michaelowa et Axel Borrmann ont notamment réalisé, en 2006, une étude comparative de cinq systèmes d'évaluation d'organismes d'aide au développement et concluent que « parmi les organisations de développement étudiées ici, seul le département d'évaluation de l'agence allemande d'aide au développement suit une procédure d'appels d'offres pour commanditer des évaluations externes. Autrement, les contacts personnels sont souvent plus déterminants que les compétences professionnelles dans la sélection des candidatures » [552]. Aujourd'hui, en 2014, la procédure des appels d'offres s'est généralisée. Lors de notre observation participante à l'AFD, nous avons constaté que la sélection des évaluateurs externes se fait dans un cadre très strict en termes de compétences et de méthodologie, réduisant le nombre de candidats potentiels. Nous constatons que les cabinets d'évaluation avec lesquels collaborent les agences d'aide sont de plus en plus nombreux. C'est tout du moins le cas pour l'AFD[553] et le DFID[554]. Ces cabinets ou évaluateurs indépendants se familiarisent vite avec les procédures d'appels d'offres, dont le contenu est très semblable d'une évaluation à l'autre, reproduisant les critères du CAD-OCDE. Selon certains évaluateurs internes aux agences d'aide, l'uniformisation des appels d'offres nuit à la réflexion et à l'innovation des évaluateurs : « *les réponses des cabinets d'évaluateurs aux*

[552] MICHAELOWA, Katharina, BORRMANN, Axel, "Evaluation Bias and Incentive Structures in Bi-and Multilateral agencies", *Review of Development economics,* vol. 10, n°6, 2006, pp. 313-329.

[553] Constat fait à partir de la base de données de l'AFD recensant tous les cabinets d'experts ayant travaillé pour l'agence dans le cadre de travaux d'évaluation. L'AFD semble travailler avec un plus grand nombre de cabinets et de consultants indépendants que le DFID. Nous pensons que cela s'explique en partie par le fait que l'AFD n'a pas seulement recours à des cabinets spécialisés dans l'évaluation mais à des bureaux d'étude d'ingénieurs sur des enjeux sectoriels du développement (eau, énergie, etc.).

[554] Constat fait à partir de la consultation des rapports d'évaluation mis en ligne par le DFID. La présence d'évaluateurs de formation ou de carrière est toutefois plus importante qu'en France.

appels d'offres sont tellement similaires d'une évaluation à une autre qu'ils font du copier-coller pour les rédiger et oublient même parfois de transformer certains termes »[555]. Pour contrecarrer cette critique, les évaluateurs externes que nous avons interrogés expliquent que « *les appels d'offres réalisés par les organismes d'aide sont souvent de qualité médiocre et le niveau d'exigence envers les évaluateurs externes est bas, ce qui ne stimule pas ces derniers à produire un travail innovant et de qualité* »[556]. Il est vrai que les appels d'offres sont très semblables, reprenant les standards du CAD-OCDE, et qu'ils peuvent être assez évasifs sur les objectifs des évaluations et les caractéristiques des projets. En effet, si les évaluations sont réalisées à la demande des départements opérationnels, la rédaction des appels d'offres ne mobilise pas toujours les acteurs de ces départements ayant dirigé les activités évaluées. Ce constat est moins vrai pour le DFID que pour l'AFD, où la communication autour des exercices des évaluations se fait plus facilement.

En France et dans une moindre mesure au Royaume-Uni, les cabinets d'expertise recrutés ne sont souvent pas spécialisés dans l'évaluation, mais dans un secteur spécifique de l'aide au développement (agronomie, grandes infrastructures, santé, etc.)[557]. Les cabinets, spécialisés dans l'évaluation des politiques publiques, existent pourtant, mais ils travaillent peu dans le secteur de l'aide au développement. Les experts engagés en tant qu'évaluateurs travaillent donc également pour les organismes d'aide sur d'autres expertises que l'évaluation, par exemple la réalisation d'études de faisabilité. Il arrive ainsi qu'un cabinet ayant réalisé des études de faisabilité sur un programme d'aide soit également recruté pour l'évaluer[558]. Si cela garantit une bonne connaissance du programme par les évaluateurs, il est aussi probable qu'ils éprouvent davantage de difficultés à y porter un regard critique, dans la mesure où ils ont participé aux études préparatoires qui ont conditionné sa mise en œuvre et ses objectifs. Les évaluateurs justifient cette pratique par le fait que « *très peu de cabinets sont compétents pour assurer des évaluations de qualité* »[559]. Selon une étude du CAD-OCDE, 50 % des évaluateurs des agences d'aide de ses pays membres estiment que les compétences des consultants externes sont faibles ou variables, alors que seulement 6 % les jugent excellentes[560]. Le recours à des experts peu qualifiés en évaluation pourrait expliquer la faible qualité de certains rapports relevée par les évaluateurs internes. De plus, si les évaluateurs sont incités par le CAD-OCDE

[555] Entretien avec une évaluatrice de la Banque mondiale, 24 octobre 2013.
[556] Entretien avec un directeur d'un cabinet d'évaluation des politiques publiques basé en France, 13 janvier 2014.
[557] La même méthode d'analyse de la base de données des consultants en évaluation de l'AFD et des rapports d'évaluation du CAD a été utilisée pour aboutir à ce constat.
[558] Ce cas de figure fut notamment observé lors de notre observation participante à l'AFD, où nous retrouvions les mêmes cabinets d'experts dans les travaux d'études préparatoires aux programmes, d'évaluation, de recherche, etc.
[559] Entretien avec un ancien directeur de l'unité de l'évaluation de l'AFD, 17 septembre 2013.
[560] OCDE, 2010 (b), *op. cit.* p. 33.

à lancer des appels d'offres à l'échelle mondiale, il est très rare qu'ils dépassent un niveau de diffusion national ou au mieux régional pour les États de l'Union européenne. Nous en concluons qu'il serait excessif de dire qu'il existe aujourd'hui un marché mondial de l'évaluation[561].

La médiocre qualité de l'expertise et des rapports des évaluateurs externes est cependant relativisée par certains évaluateurs internes aux agences d'aide qui l'expliquent par « *la trop grande complexité des méthodologies demandées, par le périmètre de plus en plus large des évaluations et par les attentes excessives des décideurs politiques sur les résultats que peuvent fournir les évaluations* »[562]. Ainsi, à des appels d'offres souvent élusifs correspondent des attentes élevées des agences d'aide quant à la qualité des rapports d'évaluation, et notamment à leur utilité directe pour la prise de décision.

Le troisième type d'indépendance est l'indépendance des évaluateurs par rapport aux influences externes, notamment dans la planification des programmes d'évaluation, la détermination du budget des unités d'évaluation et le recrutement des experts externes. Or, dans la plupart des agences d'aide bilatérales, les directeurs des unités d'évaluation négocient leur budget annuel auprès de la direction suivant les mêmes procédures que les autres départements. Le budget des unités d'évaluation est donc conditionné par l'intérêt que porte la direction à cette fonction. Les structures dirigeantes des agences d'aide ont également un droit de regard sur les programmes annuels d'évaluation. La plupart des départements d'évaluation élaborent leurs programmes d'évaluation en consultation avec les départements opérationnels et stratégiques et avec les agences locales dans les pays bénéficiaires. Selon un ancien directeur de la division de l'évaluation de l'AFD, « *cette programmation consultative des évaluations permet de s'assurer que ces dernières répondent bien aux besoins des autres départements de l'agence. Les opérationnels sont alors plus volontaires pour s'impliquer dans les évaluations, ce qui favorise leur utilisation et l'apprentissage en interne* »[563]. Le fait que les évaluateurs soient influencés par les opérateurs et les dirigeants dans le choix des évaluations serait donc une bonne chose, même si cela contrevient au principe d'indépendance. Ainsi, à la DANIDA,

[561] Cette hypothèse sera approfondie dans le chapitre 6 de la thèse.
[562] Propos recueillis auprès d'un évaluateur de l'AFD pendant une conversation informelle lors de l'observation participante réalisée en 2012-2013.
[563] Laurent Fontaine, allocution à la Société française d'évaluation, novembre 2012.

70 % des évaluations sont réalisées à la demande des autres départements[564], à l'AFD ce taux s'élève à 80 %[565].

Une autre tendance, commune à l'ensemble des organismes d'aide au développement, qui a conduit les évaluateurs à considérer qu'on les avait « dépossédés » de leur travail, est la décentralisation des activités d'évaluation[566]. Le CAD-OCDE avait diagnostiqué que si les recommandations des évaluations n'étaient pas bien appropriées par les acteurs des agences d'aide, c'était parce que leur mise en œuvre était trop centralisée[567]. Il préconisait que cette structure centralisée empêcherait une diffusion de la culture de l'évaluation dans les agences d'aide. La décentralisation des évaluations de projets devait permettre de multiplier le nombre d'évaluations réalisées et de faciliter la diffusion de leurs conclusions au sein des agences d'aide. Les organismes d'aide seraient ainsi moins bureaucratiques et plus capables d'apprendre et de s'adapter aux recommandations des évaluations. Selon Michel Crozier, en effet, un système bureaucratique se caractérise par un, « circuit « erreurs-informations-corrections » qui fonctionne mal et où il ne peut y avoir de correction rapide des programmes d'action »[568]. En améliorant la circulation des informations produites par les évaluations, on réduirait ainsi la lourdeur bureaucratique des agences d'aide.

Avec la décentralisation des évaluations, les évaluateurs regrettent d'être réduits à une tâche peu stimulante de supervision des agences de terrain[569]. Cependant, il ne faut pas oublier que la décentralisation a aussi répondu à une nécessité pratique pour les organismes d'aide. En effet, l'exigence d'évaluation des tutelles des agences d'aide, mais aussi des dirigeants de ces organismes, se tournant vers des travaux de plus en plus complexes, analysant l'efficacité des programmes sectoriels ou de l'ensemble des activités dans un pays, il devenait nécessaire de déléguer une partie de leur travail - les évaluations de projets, les plus simples à réaliser - aux agences de terrain. Alors, le personnel des agences délocalisées gère la mise en œuvre des évaluations de projets. Cette gestion par le personnel de terrain est aussi censée permettre un meilleur contact avec les parties prenantes des projets et les consultants en évaluation locaux, ainsi qu'une amélioration de la diffusion d'une culture de l'évaluation dans les agences décentralisées. Les évaluateurs du siège remettent en cause ces effets présumés : « *la culture de*

[564] SCHAUMBÜRG, Miller, "Use of aid evaluation from an organizational perspective", *Evaluation,* vol. 11, n°2, pp. 207-222.
[565] Laurent Fontaine, allocution à la Société française d'évaluation, novembre 2012.
[566] Pour rappel, une répartition des tâches s'est organisée dans les années 2000 entre le département de l'évaluation qui se consacre aux évaluations de grande échelle (évaluations de programme, sectorielles, stratégique, par pays, d'instruments de développement, etc.) et les agences de terrain qui gèrent les évaluations de projets.
[567] DAC-OECD, 2006, *op. cit.*
[568] CROZIER, Michel, 1963, *op cit*. p. 247.
[569] Entretien avec un évaluateur de l'AFD, 24 juin 2013.

l'évaluation n'a pas encore véritablement imprégnée les agences décentralisées, où le personnel s'implique peu dans l'évaluation des projets externalisées à des experts externes »[570].

Selon nous, cette affirmation est à demi fondée. L'influence positive de la décentralisation sur la diffusion d'une culture de l'évaluation au personnel de terrain est très inégale d'une agence d'aide à l'autre, ainsi qu'au sein d'une même agence. Elle dépend grandement de la sensibilité de chacun des acteurs vis-à-vis de l'évaluation, qu'ils vont parfois voir comme une opportunité d'apprentissage, mais parfois aussi comme une obligation administrative à remplir[571]. De manière générale, le personnel des agences de terrain de l'AFD considère les évaluations comme utiles, mais critique leur systématisation qui en fait une tâche trop lourde. Le nombre d'évaluations empêcherait d'en tirer des leçons car le temps de la réflexion et du recul s'en retrouverait réduit[572].

Nous observons, à travers ces trois transformations des systèmes d'évaluation des agences d'aide (indépendance organisationnelle des unités centrales d'évaluation, externalisation des évaluations de grande échelle et décentralisation des évaluations de projets), que les dirigeants des agences d'aide et les départements d'évaluation ont abordé l'enjeu de la diffusion d'une culture de l'évaluation par l'instrument des réformes organisationnelles. Ces réformes sont souvent mal vécues par les évaluateurs « seniors » qui considèrent perdre de leur pouvoir de changement dans les organismes d'aide à mesure que l'évaluation est externalisée et décentralisée. L'un d'entre eux témoigne : « *l'externalisation et la décentralisation des évaluations ont respectivement eu lieu en 2006 et 2009. Avant, elles étaient internalisées et centralisées et les évaluateurs internes produisaient les évaluations en collaboration avec des consultants externes, parfois du Sud. C'était ce travail qui était le plus intéressant. Maintenant, toutes les évaluations sont externalisées et il y a une perte de contrôle sur leur contenu* »[573].

Nous allons désormais démontrer que ces réformes techniques et organisationnelles sont insuffisantes pour entraîner l'émergence d'une culture de l'évaluation dans des organisations dont la culture ne se prête pas à l'apprentissage au sens large. Nous considérons que des évaluations mises en œuvre de manière internalisée, en consultation étroite avec les autres départements des agences, et centralisée peuvent rester indépendantes si des mécanismes de

[570] Entretien avec un ancien directeur du département d'évaluation de l'AFD, 17 septembre 2013.
[571] Entretien avec un chargé de projets de l'AFD au Laos, 30 juin 2012 ; Entretien avec un chargé de projet pour une agence d'aide au développement au Sénégal, 3 avril 2014.
[572] Propos recueillis lors de notre séjour au sein de l'agence de l'AFD au Laos en juin et juillet 2012 ainsi que lors de rencontres avec les représentants d'autres agences délocalisées lors de notre observation participante au siège, voir liste des personnes interviewés au Laos en annexes 8 et 9.
[573] Entretien avec un évaluateur de l'AFD, 14 juin 2013.

contrôle sont mis en place et que les évaluateurs ont intégré une éthique professionnelle forte (comme c'est le cas des auditeurs internes). Dans ce système, il serait également possible que des évaluateurs externes collaborent avec les évaluateurs internes pour mener à bien les travaux d'évaluation. Un tel système permettrait de mieux conjuguer redevabilité externe et apprentissage interne dans la fonction d'évaluation. Toutefois, sa mise en place demanderait un travail sur la culture organisationnelle des agences d'aide, qui a été mis de côté au profit de réformes institutionnelles plus visibles et plus rapides à instaurer. Nous souhaitons mettre en avant les liens entre la diffusion d'une culture de l'évaluation et la préexistence d'une culture de l'apprentissage chez les acteurs des agences d'aide ; ces deux aspects de la culture organisationnelle se renforçant mutuellement.

- **Culture de l'évaluation et organisations « apprenantes »**

Pour qu'une organisation parvienne à adopter une culture de l'évaluation, c'est-à-dire à faire que l'évaluation soit une fonction dont les pratiques et les structures sont légitimées, dont les valeurs sont partagées et dont les messages politiques sont pris en compte de manière consensuelle ou non[574], plusieurs facteurs rentrent en compte. Ces facteurs ont été énoncés dans une étude originale d'Isabelle Bourgeois et Bradley Cousins, deux évaluateurs canadiens[575].

Le premier facteur est « l'environnement externe », soit dans notre cas, le contexte international et national des organismes d'aide, dont nous avons traité dans la première partie de cette recherche. Il est ressorti de notre analyse que les organismes d'aide évoluent aujourd'hui dans un contexte en demande d'évaluation. Toutefois, un type spécifique d'évaluation a tendance à être favorisé : les évaluations expérimentales à des fins de redevabilité.

La structure organisationnelle de l'évaluation (deuxième facteur), étudiée au début de ce chapitre, joue un rôle certain sur la diffusion d'une culture de l'évaluation chez les acteurs des organismes d'aide. Toutefois, nous développons l'idée que la diffusion d'une culture de l'évaluation découle surtout de la culture de ces organismes. Nous allons donc nous y attarder un peu plus avant de passer au quatrième et dernier déterminant (le soutien de la direction des organisations).

[574] Cette description de ce que nous considérons être une organisation possédant une culture de l'évaluation est inspirée de la définition du terme "culture" du Dictionnaire de la science politique de Hermet et al, voir introduction.
[575] BOURGEOIS, Isabelle, COUSINS, Bradley. J, "Understanding Dimensions of Organizational Evaluation Capacity" *American Journal of Evaluation*, vol. 34, n°3, 2013, pp. 299-319.

Nous défendons l'idée qu'une culture de l'évaluation ne peut se développer que dans une organisation qui encourage le questionnement, l'expérimentation et l'apprentissage par un processus d'essai-erreur. Il faut également que les acteurs de l'organisation soient incités à partager leurs expériences et compétences, à se former, ainsi qu'à porter un regard critique sur leur travail. Par ailleurs, il faut que le personnel chargé de la production de connaissance soit formé à la gestion de cette connaissance. Lorsque nous avons interrogé des évaluateurs sur ce qu'est, selon eux, une culture de l'évaluation, ils répondent par exemple : « *la culture de l'évaluation inclut une volonté de tester la réalité, de se poser la question : comment savons-nous ce que nous pensons connaître ? Elle demande d'utiliser les données des évaluations pour nourrir la décision. La culture de l'évaluation n'est pas limitée à la réalisation d'évaluations formelles. C'est une manière de penser analytique qui doit imprégner toutes les activités de l'organisation* »[576]. Ou encore, « *la culture de l'évaluation, c'est la culture de la remise en question de nos idées préconçues, de l'autocritique, c'est la culture du dialogue collectif et de l'apprentissage à partir des travaux d'évaluation* »[577]. Dans ces citations transparaît le lien entre culture de l'apprentissage et culture de l'évaluation.

Selon Peter Senge, les « organisations apprenantes »[578], c'est-à-dire qui ont une culture de l'apprentissage sont « des organisations où les individus étendent continuellement leur capacité à atteindre les résultats qu'ils désirent, où les nouveaux modes de pensées sont encouragés, où l'aspiration collective est libérée et où les individus apprennent continuellement ensemble »[579]. On retrouve donc de nombreux points communs entre cette définition de la culture de l'évaluation et celle donnée par nos interviewés. Il est frappant de voir que l'évaluation est toujours perçue comme un instrument d'apprentissage par les évaluateurs et non comme un élément de redevabilité, notion qui ressort peu des définitions données durant les entretiens que nous avons menés. Ainsi, il y a une tension dans l'objet de l'évaluation entre l'usage gestionnaire qu'en font les dirigeants des agences d'aide et la fonction « transformatrice » que lui donnent les évaluateurs.

Nous retrouvons ce rapprochement entre organisation apprenante et culture de l'évaluation dans un article de Herink Schaumburg-Müller datant de 2005 : « dans une organisation apprenante, il est important de réaliser des évaluations qui fournissent des leçons

[576] Entretien avec un évaluateur des Nations unies, 6 octobre 2014.
[577] Entretien avec un consultant en évaluation des politiques publiques, 12 janvier 2013.
[578] Peter Senge et Chris Argyris parlent « d'organisations apprenantes », concept qui sert à étudier la manière dont les organisations apprennent et quelles sont les conditions pour qu'elles deviennent « apprenantes », c'est-à-dire des organisations qui facilitent l'apprentissage de leurs membres et qui se transforment continuellement à partir de cet apprentissage. C'est cette capacité d'adaptation qui assure la survie de l'organisation dans son environnement compétitif.
[579] SENGE, Peter, *The fifth discipline: the art and practice of the learning organization,* Doubleday/Currency, 1990, 424p.

à partir desquelles l'organisation peut apprendre, innover et se développer. L'organisation apprenante est le modèle de l'organisation moderne, opposée à l'organisation traditionnelle bureaucratique, hiérarchique, autocratique et centralisée »[580].

Cette littérature et les entretiens que nous avons menés laissent penser que l'apprentissage émerge des interactions sociales. L'émancipation du personnel est alors vue comme cruciale, ainsi que le dialogue transversal et le travail collaboratif. Selon Chris Argyris et Donald Schön, étudier le processus d'apprentissage d'une organisation revient à analyser comment ses membres arrivent à établir une compréhension commune de problèmes particuliers et à forger des solutions appropriées pour y remédier[581]. C'est-à-dire leur capacité à construire une culture politique commune qui comprend la solution de l'évaluation pour investiguer et résoudre des problèmes particuliers. En effet, selon Gabriel Armond et Sydney Verva, l'une des dimensions de la culture politique est « évaluative » et désigne la manière dont les acteurs d'un système politique vont produire un jugement commun de ce système et proposer des solutions communes pour l'améliorer ou résoudre les problèmes qui se posent[582].

D'après Chris Argyris et Donald Schön, une compréhension commune est difficile à atteindre dans des organisations où les critères de réussite sont ambigus, où les membres ont des intérêts divergents et où le pouvoir est inégalement distribué[583]. Ce sont des caractéristiques présentes dans les agences d'aide, comme en témoigne Henrik Schaumburg-Müller : « les agences d'aide ont besoin d'apprendre et d'innover, mais elles sont aussi des organisations bureaucratiques et fragmentées avec des structures hiérarchiques qui ont été établies avec d'autres objectifs que l'apprentissage »[584]. Transparaissent ici les difficultés de diffusion de la connaissance et de l'expérience au sein des agences d'aide. En effet, certains agents de l'AFD pointent la rigidité de la structure hiérarchique de l'agence comme un obstacle majeur aux échanges d'idées et à l'apprentissage[585].

Il est également vrai que les objectifs que ces agences d'aide poursuivent sont ambigus et multiples, non priorisés et parfois contradictoires les uns avec les autres. Pour illustrer ce constat, nous pouvons, par exemple, citer l'économiste William Easterly : « la contradiction

[580] SCHAUMBURG-MULLER, Henrik, 2005, *op. cit.*
[581] ARGYRIS, Chris, SCHÖN, Donald. A, *Organizational learning II: Theory, Method and Practice*, Reading, MA: Addison-Wesley, 1996, 305p, p. 6.
[582] ALMOND Gabriel, VERBA, Sydney, *The Civic Culture, Political Attitudes and Democracy in Five nations*, Princeton University Press, 1963.
[583] ARGYRIS, Chris, SCHÖN, Donald, 1996, *op. cit.*
[584] SCHAUMBÜRG-MUELLER, Henrich, *op. cit.*
[585] Propos issus d'une discussion informelle avec des oéprationnels de l'AFD lors de notre observation participante en 2012 et 2013.

entre l'objectif d'appropriation de l'aide au développement par les bénéficiaires et les Objectifs du Millénaire établis arbitrairement (par les donateurs) reste non remarquée. La contradiction entre la participation des pays bénéficiaires et l'imposition de *Poverty Reduction Strategy Plan* uniformisés par les banques de développement également. (…) Dans les agences d'aide, il n'y a pas de hiérarchisation, ni de sélection entre ces objectifs qui entrent en compétition et en contradiction »[586]. Cet état de fait rend difficile aussi bien l'apprentissage que l'évaluation. Dans ce contexte, il n'est pas étonnant de voir que les évaluateurs ne puissent guider et conseiller les acteurs du développement qui considèrent, dès lors, l'évaluation comme inutile.

L'apprentissage interne et le rôle de l'évaluation dans ce processus ne font l'objet de travaux de réflexion par les organismes d'aide que depuis le début des années 2000, moment où leur incapacité à apprendre de leurs expériences passées a été largement critiquée par le monde académique[587]. Un des chercheurs dont les critiques ont été les plus retentissantes est William Easterly qui dénonce l'illusion des rapports d'évaluation « brillants » des organisations d'aide au développement, notamment de la Banque mondiale, et la publication de nombreuses *success stories* qui cacheraient une absence d'apprentissage : « la communauté de l'aide ne veut pas porter un regard critique sur le passé parce qu'une telle initiative pourrait entraîner une publicité négative et menacer sa capacité à lever de nouveaux fonds. (…). La communauté de l'aide s'organise comme un 'cartel de bonnes intentions' supprimant les retours critiques et l'apprentissage à partir du passé et l'identification de nouvelles pratiques pour améliorer l'efficacité de l'aide »[588].

Les évaluateurs que nous avons interrogés mettent souvent en avant la faiblesse de la culture de l'apprentissage dans leurs organisations pour expliquer les difficultés qu'ils éprouvent à faire valoir les résultats de leur expertise[589]. Si nous prenons l'exemple de la Banque mondiale, aujourd'hui pourtant considérée comme un modèle d'organisation apprenante, les témoignages d'évaluateurs des années 1970 à 2000 ont souligné la faiblesse de l'apprentissage interne et les conséquences que cela a eu sur leur fonction d'évaluateurs[590]. Christopher Willoghby, évaluateur dans les années 1970, déclare par exemple : « notre principal

[586] EASTERLY, William, *The cartel of good intentions: the problem of bureaucracy in foreign aid*, Center for Global Development, octobre 2002, p. 48.
[587] SIDA, *knowledge and learning in aid organizations: a literature review with suggestions for further studies*, SADEV working paper, janvier 2007; Independent Commission for Aid Impact, *How DFID learns*, report 34, 2014; JONES, Harry, MENDIZABAL, Enrique, *Strengthening learning from research and evaluation: going with the grain*, ODI, septembre 2010.
[588] EASTERLY, 2002, *op. cit.*
[589] Sentiment qui apparaît généralisé au sein de l'AFD ; entretien avec une analyste des politiques de l'unité d'évaluation du CAD-OCDE, 2 juillet 2014 ; entretien avec une chargée de projets en évaluation au F3E, 19 avril 2012
[590] Archives du séminaire organisé en 2000 par la Banque mondiale à l'occasion du 30ème anniversaire de son département d'évaluation.

espoir était de réussir à établir une tradition de l'autocritique et de la réflexion sur les opérations passées, sans que ces exercices soient empêchés par la peur de la confrontation à des résultats médiocres, mais au contraire soient motivés par l'envie d'apprendre des leçons du passé. Nous (les évaluateurs) attendions le jour où il y aurait une meilleure mémoire institutionnelle au sein de la Banque qui génèrerait davantage d'informations sur la manière dont la Banque dépense ses ressources sur les différents aspects des projets »[591].

Pour beaucoup, ce jour est finalement venu à la fin des années 1990, sous l'impulsion de James Wolfensohn, alors président de la Banque mondiale, qui déclara dans son discours de prise de fonction en 1996 vouloir faire de la Banque mondiale une « banque de connaissances »[592]. Au-delà d'améliorer la redevabilité, l'efficience et l'efficacité des activités de la Banque, l'objectif de James Wolfensohn était de la positionner comme la principale source de production et de diffusion d'expertise sur le développement. La Banque mondiale fut ainsi la première organisation de développement à faire de l'apprentissage interne un axe de stratégie politique. Cette stratégie a été considérée comme une réussite par les donateurs. En conséquence, les plus petites agences d'aide bilatérales se sont rangées derrière l'expertise de la Banque mondiale, réduisant leur expertise en interne, estimant que la « Banque pouvait conduire la réflexion sur le développement pour eux »[593]. La Banque mondiale parvint également à évincer la voix des autres organisations multilatérales. Elle plaça un grand nombre de ses experts à des postes stratégiques des Nations unies. Le directeur actuel du PNUD, par exemple, est un ancien vice-président de la Banque mondiale. L'influence de la Banque a également été grandissante au sein de l'OCDE. Ce cas illustre bien les enjeux de pouvoir sous-jacents aux dimensions d'apprentissage, d'évaluation et d'expertise pour les organisations internationales. En renforçant son apprentissage interne, sa pratique de l'évaluation et ainsi l'accumulation de connaissances, la Banque mondiale a renforcé son mandat et son statut de référence internationale d'expertise sur le développement. Ainsi, sa légitimité pour imposer sa vision politique du développement aux autres acteurs a été renforcée.

La stratégie de faire de la Banque mondiale « une banque de connaissance » a donc contribué à la montée en puissance de l'organisme dans le système international. Devant le succès de la stratégie de la Banque mondiale, les autres organismes d'aide ont, dans un second temps, commencé à mettre en place des systèmes de gestion interne de la connaissance. Les

[591] Archives du séminaire organisé par la Banque mondiale en 2000 à l'occasion du 30ème anniversaire de son département d'évaluation.
[592] WOLFENSOHN, James, *People and Development. Annual Meeting Address,* World Bank, octobre 1996.
[593] Entretien avec un ancien évaluateur du DFID, 17 juin 2013.

études commanditées ou menées par les agences d'aide sur cette thématique font beaucoup référence à la Banque mondiale. Nous retrouvons communément dans ces travaux des affirmations, telles que « l'expérience de la Banque mondiale en termes d'apprentissage interne a eu une influence importante sur notre politique de gestion de la connaissance »[594]. Nous observons, une fois encore, le rôle que jouent les organisations multilatérales dans la production de bonnes pratiques de gestion de l'aide au développement qui vont se diffuser internationalement par le mimétisme des agences d'aide soucieuses d'être perçues comme des organisations modernes et efficaces.

Toutefois, la manière dont l'expérience de la Banque a été interprétée par les acteurs des différentes agences varie significativement. Alors que les membres du DFID et de la JICA élaborent des stratégies de gestion de la connaissance largement influencées par la Banque mondiale, les agents de la SIDA ont une vision plus sceptique des pratiques de la Banque et développent un mode plus informel de gestion de la connaissance[595]. Nous voyons dans ces différences le reflet de conceptions disparates des modalités d'apprentissage dans les politiques publiques nationales, ainsi que dans la culture organisationnelle des agences. Dans certaines agences d'aide, le partage des connaissances et l'apprentissage sont perçus comme se réalisant par des interactions informelles et non structurées et par le dialogue quotidien, au sein des différents groupes de professionnels qui apprennent les uns des autres (SIDA/NORAD). Cette approche est la plus répandue dans les pays du Nord de l'Europe qui ont une longue tradition de la transparence et de la diffusion de l'information dans leurs politiques publiques. Pour d'autres, la diffusion de la connaissance nécessite la mise en place d'un encadrement très formalisé des interactions (via des procédures diverses) qui vont mener à l'apprentissage (DFID, Banque mondiale). C'est souvent le cas dans les États où la diffusion de l'information et de la connaissance au sein des administrations publiques est très formalisée et encadrée[596].

Quelles que soient les modalités de gestion de la connaissance, les évaluateurs des agences bilatérales sont très critiques quant à leur efficacité[597]. Par exemple, un évaluateur de la SIDA explique que « bien que le *Management Response System* apparaisse bien conçu en théorie, il ne parvient pas à atteindre son objectif d'amélioration de l'apprentissage organisationnel. Les raisons de ce manque d'efficacité incluent sa faible vitesse de mise en

[594] SIDA, 2007, *op. cit.*
[595] *Ibid*.
[596] KING, Kenneth, Mc GRATH, Simon, *Knowledge sharing in development agencies: lessons from four cases*, the World Bank operations evaluation department, 2003.
[597] FORESTI, Maria, 2007, *op. cit.*

œuvre et sa faible intégration aux enceintes de prise de décision »[598]. Les évaluateurs du DFID, quant à eux, mettent en avant le manque de communication et de coordination internes entre le département de l'évaluation et les autres départements de l'agence[599].

La Banque mondiale est vue comme faisant figure d'exception dans sa réussite à développer une culture de l'apprentissage. Toutefois, les témoignages que nous avons recueillis d'évaluateurs actuels de la Banque mondiale à ce sujet se contredisent. Certains soulignent que les managers de programmes sont fortement incités à prendre en compte les leçons tirées d'expériences passées, à s'autoévaluer et à intégrer toutes les parties prenantes des programmes dans les processus d'évaluation. Toutefois, certains évaluateurs sont plus mitigés quant à l'utilisation des évaluations par le personnel de la Banque mondiale et considèrent que la capacité de la Banque à apprendre des évaluations qu'elle mène est surévaluée. Une évaluatrice raconte que « *les résultats des évaluations ne sont pas pris en compte dans l'élaboration des nouveaux projets. Il n'y a pas d'incitations pour que les managers de projets les utilisent comme base de départ pour planifier les nouveaux projets qui sont dans la continuité de ceux ayant déjà été évalués. De plus, il est de l'intérêt de toutes les parties prenantes, Banque mondiale et ses partenaires du Sud, de fixer des objectifs faciles à atteindre qui contournent les difficultés observées par les évaluations* »[600]. Pour vérifier cette affirmation, nous avons consulté une étude, réalisée par une évaluatrice de la Banque mondiale, analysant les documents de planification de projets reconduits par la Banque à la suite à d'évaluations. Nous avons constaté qu'ils reprennent très marginalement les conclusions et les recommandations des évaluations[601].

Ces témoignages nous montrent que la culture de l'apprentissage est très variable au sein de la Banque mondiale et dépend surtout de la volonté individuelle des acteurs à prendre en compte les évaluations dans leur travail. Si la Banque mondiale est une « banque de savoir », La diffusion de ce savoir n'est donc pas encore effective. La mise en place de mécanismes de diffusion de la connaissance ne suffit pas à assurer la construction d'une culture de l'apprentissage généralisée dans les agences d'aide. Il est nécessaire de travailler également auprès des acteurs pour les sensibiliser à l'importance de l'apprentissage interne et à l'utilité de l'évaluation pour cet apprentissage.

[598] SIDA, 2007, *op. cit.*
[599] DIFD, *Doing the knowledge: how DFID compares with best practices in knowledge management*, 2000.
[600] Entretien avec une évaluatrice de la Banque mondiale, 20 décembre 2013.
[601] Étude menée par une évaluatrice de la Banque mondiale sur une vingtaine de projet portant d'urbanisation et de formation professionnelle en Afrique subsaharienne.

Par ailleurs, nous notons que la Banque mondiale s'est attachée à promouvoir une image d'organisation apprenante seulement lorsque la société civile et ses financeurs ont commencé à émettre des critiques fortes sur son incapacité à tirer des leçons des expériences passées et sur les impacts négatifs des programmes qu'elle finançait, notamment en matière d'environnement, donc à un moment où sa puissance symbolique était affaiblie. C'est donc dans une logique de maintien de sa légitimité et de son mandat que l'organisation a adopté une culture de l'apprentissage nécessitant la remise en cause de ses préconceptions en matière de bonnes pratiques de développement[602]. La Banque a ainsi trouvé un moyen de se redéployer, de se relégitimer, en adoptant des bonnes pratiques de gestion de la connaissance. Les agences bilatérales, moins soumises à la critique de la société civile internationale, car moins visibles, n'ont pas ressenti ce même besoin de légitimation. Leur motivation à développer des mécanismes de gestion de la connaissance s'explique principalement par l'influence des organisations multilatérales, en premier lieu desquelles la Banque mondiale, qui ont diffusé des bonnes pratiques en la matière. La réputation des agences bilatérales, n'est néanmoins pas conditionnée en premier lieu par l'adoption de ces réformes de gestion de la connaissance. En conséquence, elles sont restées des enjeux secondaires parmi les nombreuses bonnes pratiques diffusées par les organismes multilatéraux[603].

Nous allons désormais montrer que de multiples facteurs rendent difficile la diffusion de la connaissance issue des évaluations et la construction d'une mémoire organisationnelle dans les agences d'aide. Premièrement, nous identifions la rotation rapide des postes, le manque de mécanismes de transmission des connaissances entre les anciens et les nouveaux agents et la proportion croissante de personnel en contrat à durée déterminée.

Ensuite, les évaluateurs de la SIDA mettent en avant la critique, déjà observée ci-dessus, du manque de compétence des évaluateurs externes : « les signaux envoyés par les évaluations sont faibles, ainsi que la qualité des évaluations. Les évaluations manquent souvent d'un point de vue pragmatique, que ce soit au niveau des projets ou des programmes. De plus, la profondeur de l'analyse est sacrifiée en faveur d'une tendance à répondre à trop de questions dans un temps trop court. En conséquence, les recommandations sont souvent vagues ou ambiguës et difficiles à mettre en pratique. Deuxièmement, il y a peu d'exemples d'évaluations qui ont un apport original en termes de connaissance. Aussi, les résultats des évaluations

[602] Raisonnement inspiré de la thèse développé par Steven R. Brechin dans l'ouvrage *Planting Trees in the Developing World : a sociology of international organizations,* où il défend l'idée que les organisations internationales ne remettent en cause leurs croyances et leurs bonnes pratiques que lorsque leur « survie » est menacée.
[603] KLEIN, Amsara, LAPORTE, Camille, SAIGET, Marie, 2015 (à paraître), *op. cit.*

apparaissent fréquemment répondre à un autre objectif que l'amélioration de l'efficacité des programmes d'aide et l'apprentissage. L'utilisation des évaluations pour légitimer des décisions déjà prises est quelque chose d'habituel. Enfin, les évaluations n'intéressent qu'une proportion très limitée des acteurs engagés dans un projet et affectés par ses résultats »[604].

Selon nous, plus que le manque de compétence des évaluateurs externes, c'est ce désintérêt du personnel des agences d'aide, ou plutôt le manque de temps dont il dispose pour participer et s'intéresser aux travaux d'évaluation qui est la clé du problème. Cette situation est tout d'abord la conséquence de la pression à décaisser des fonds, dont font l'objet les managers des organismes d'aide. Ce « désintérêt » entraîne une demande d'évaluation floue, un suivi des travaux d'évaluation intermittent par les comités de pilotage et une prise en compte aléatoire des recommandations des évaluations. Ce faible suivi des travaux d'évaluation par les acteurs des agences d'aide, qu'ils soient opérationnels, décideurs, ou même parfois évaluateurs, fait que les évaluateurs externes se déclarent souvent « déçus » de leur partenariat avec les organismes d'aide[605].

Ce n'est donc pas la quantité d'informations, notamment issue des évaluations, qui pose problème à l'apprentissage, mais la difficulté à impliquer le personnel des agences d'aide aux évaluations, les deux problèmes pouvant être liés. Le personnel des agences d'aide fait même face à une information trop abondante qu'il devient difficile de synthétiser et d'intégrer[606]. Par ailleurs, nous comprenons, à travers plusieurs récits d'expérience[607] que l'adoption de nouvelles pratiques et la réforme des pratiques existantes sont souvent jugées plus risquées que profitables par le personnel des agences d'aide. Cela peut être dû au fait qu'il existe peu d'incitations pour encourager le personnel à innover dans ces organismes. Ce manque d'incitations rend difficile l'apprentissage à partir des évaluations, qui nécessite une transformation des pratiques existantes et la mise en œuvre d'expérimentations qui peuvent être risquées. Par exemple, à l'AFD, pour améliorer l'utilisation interne des évaluations à des fins d'apprentissage, la division de l'évaluation a été placée dans le département *Recherche et capitalisation*. Toutefois, les responsables des unités d'évaluation et de recherche reconnaissent que les liens organisationnels entre les divisions produisant de la connaissance et les départements

[604] SIDA, *Are SIDA's evaluations good enough? An assessment of 34 evaluations report,* Sida Studies in Evaluation, janvier 2008.
[605] Entretien avec un consultant en évaluation des politiques publiques, 17 septembre 2014 ; entretien avec le directeur d'un cabinet d'évaluation des politiques publiques basé en France, 9 décembre 2014 ; entretien avec le directeur d'un cabinet d'évaluation des politiques publique basé en France, 13 janvier 2014.
[606] SIDA, 2007, *op. cit.*
[607] SIDA, 2008, *op. cit.*
PRITCHETT, Lant, "It pays to be ignorant: A simple political economy of rigorous program evaluation", *the journal of policy reform,* 2002, pp. 251-269.

opérationnels sont trop faibles, ce qui empêche l'effet positif espéré par cette réforme de se produire[608]. Il manque des mesures encourageant et incitant les acteurs de différents départements à apprendre les uns des autres.

Enfin, l'externalisation des évaluations convie le risque d'un affaiblissement de l'apprentissage. En effet, la mission des évaluateurs externes s'arrête à la production du rapport. Les consultants externes n'ont souvent pas idée de l'utilisation qui sera faite de leur travail une fois qu'ils l'ont transmis aux organismes d'aide[609]. Maria Foresti, chercheuse à l'ODI au Royaume-Uni, relate que « l'implication directe du personnel dans les évaluations est de plus en plus considérée comme une opportunité pour augmenter l'apprentissage individuel et organisationnel, bien que la plupart des agences d'aide (toutes exceptées les banques de développement) s'appuient sur des consultants externes pour mener les évaluations. La consultation avec les départements opérationnels et stratégiques est plus intense dans les phases de planification de l'évaluation, de revue du rapport et de diffusion de l'évaluation »[610]. Cela est parfois vrai, toutefois, notre expérience au sein de l'AFD et nos dialogues avec des évaluateurs du DFID ont montré que les évaluateurs internes ont du mal à jouer le rôle de « courroie de transmission » entre les évaluateurs externes et le personnels des organismes d'aide et à trouver du temps pour diffuser à ces derniers les conclusions des travaux d'évaluation[611]. L'utilisation des évaluations dépend alors de la volonté de chacun de s'informer sur les rapports produits. Par ailleurs, nous estimons que l'apprentissage opère surtout par l'implication du personnel des agences durant la phase de production de l'évaluation et non en amont et en aval de l'exercice. Au-delà de l'apprentissage, c'est en réalisant des évaluations que les acteurs de l'aide pourront s'imprégner d'une culture de l'évaluation.

Enfin, nous allons nous intéresser à la quatrième et dernière condition d'une diffusion de la culture de l'évaluation dans les agences d'aide : le soutien de la direction. À l'AFD, c'est en grande partie grâce à Jean-Michel Severino[612], directeur général de l'agence de 2001 à 2010 et à Jean-David Naudet[613], responsable de l'unité de l'évaluation de 2005 à 2010, que la culture

[608] Constats établis lors de notre observation participante à l'AFD.
[609] Entretiens formels et informels menés avec des évaluateurs externes travaillant auprès d'organismes d'aide au développement et d'autres administrations publiques lors de stages et de consultances entre 2011 et 2014.
[610] FORESTI, Maria, 2007, *op. cit.*
[611] Constat issu de l'observation de réunions d'équipe entre les évaluateurs de l'AFD. Pour le DFID entretien avec une évaluatrice interne, 4 juillet 2014 et avec un évaluateur indépendant, 18 février 2014.
[612] Avant de devenir directeur général de l'AFD en 2001, Jean-Michel Sévérino a été vice-président de la Banque mondiale pour l'Asie. De ce passage à la Banque mondiale, il a compris l'importance que revêtent l'apprentissage, l'évaluation et la capitalisation de connaissances. Dimensions qu'il essaya de renforcer par la suite à l'AFD.
[613] On peut penser que le fait que Jean-David Naudet (statisticien et économiste), en plus d'être un praticien de l'aide au développement, ait des activités de recherche au sein du DIAL et du Club du Sahel, notamment sur l'efficacité de l'aide et l'évaluation, l'ait rendu plus à même et volontaire de diffuser une culture de l'évaluation en interne à l'AFD.

de l'évaluation a grandi au sein de l'agence. Le premier donnait une grande importance au développement de la recherche et de l'évaluation et au renforcement des liens entre connaissance et activités opérationnelles et stratégiques. Le second a travaillé au rapprochement entre les évaluateurs et les opérateurs, ainsi qu'à la diffusion des travaux d'évaluation auprès des autres départements de l'agence. Par contraste, Dov Zerah, directeur général de l'AFD de 2011 à 2013, n'avait pas le même intérêt pour la promotion des activités de recherche et d'évaluation. Laurent Fontaine, directeur de l'unité de l'évaluation de 2011 à 2014, était, de son propre aveu, peu intéressé par l'évaluation avant sa prise de fonction, ce qui a nui à sa capacité de faire valoir cette expertise au sein de l'agence. Ce manque de soutien a entraîné un affaiblissement de l'influence de l'évaluation à l'AFD[614]. Un évaluateur raconte : « *Jean Michel Severino, anciennement à la Banque mondiale, est quelqu'un qui, sur le plan de la recherche, a mis l'accent sur la production de connaissances pour faire un peu le parallèle avec ce que faisait la Banque mondiale à cette époque. Son idée était de ne pas seulement participer au financement du développement, mais également à la production d'une expertise. L'évaluation a été logée au département de la recherche avec une vocation première de production de connaissances. Par la suite, il y a eu une nouvelle direction, moins favorable à la recherche et à l'évaluation* »[615]. À travers l'exemple de l'AFD, nous voyons qu'au-delà des structures et des procédures, c'est avant tout l'engagement et le dynamisme des acteurs qui font naître une culture de l'évaluation et de l'apprentissage. Dans la continuation de cette observation, nous allons désormais analyser de manière plus approfondie le poids des acteurs dans la construction d'une culture de l'évaluation dans les organismes d'aide, puis étudier les relations que les évaluateurs entretiennent aujourd'hui avec les différents niveaux d'acteurs au sein de ces organismes.

B. Jeux d'acteurs et rapports de pouvoir

À travers le cas de la Banque mondiale, nous allons démontrer que les acteurs ont eu un rôle de premier ordre dans la montée en importance de l'évaluation dans les organismes d'aide durant les quarante dernières années. Les réformes structurelles qui ont été mises en place pour donner une place centrale à l'évaluation et garantir la qualité des travaux d'évaluation ont été

[614] Ce constat est tiré de notre observation participante à l'AFD de 2012 à 2014, où nous avons pu interroger de nombreux membres de l'organisme sur le rapport à l'évaluation des différents niveaux hiérarchiques de l'agence.
[615] Entretien avec un évaluateur de l'AFD spécialiste des méthodes expérimentales, 19 novembre 2013.

permises par des personnalités qui ont porté et incarné des approches nouvelles de la gestion de l'aide au développement. Notre ambition est de faire ressortir le rôle qu'ont joué ces quelques acteurs clés, caché derrière l'histoire institutionnelle désincarnée de la diffusion d'une culture de l'évaluation dans l'organisation. Nous avons fait le choix de la Banque mondiale, car ses membres ont impulsé une culture de l'évaluation bien au-delà de l'organisation sur la scène internationale. De plus, la Banque mondiale a une politique de l'évaluation très développée par rapport aux autres banques de développement et agences d'aide.

- **Le poids des acteurs dans la construction d'une culture de l'évaluation : le cas de la Banque mondiale**

En 1971, c'est le président de la Banque, Robert Mc Namara qui a donné de l'importance à la fonction d'évaluation. Selon Mervin Weiner, ancien évaluateur à la Banque mondiale, « l'évaluation était une fonction qui faisait partie de la conception du bon management selon Robert Mc Namara. Celui-ci a donc eu l'ambition de mettre en place une fonction d'évaluation au sein de la Banque dès le début de son mandat. La légende veut que l'évènement déclencheur de cette décision ait eu lieu lors d'une réunion du conseil d'administration, alors que la Banque venait juste de commencer à financer des initiatives de développement agricole et rural. Un directeur exécutif aurait alors demandé par quel moyen la Banque pourrait savoir comment ces projets évolueraient après leur mise en œuvre »[616]. La réponse de Robert Mc Namara aurait été « par l'évaluation ». Son espoir était de « *construire un système d'évaluation qui agirait comme une grille de notation pour mesurer les accomplissements de la Banque mondiale, fournir des retours plus approfondis sur les projets afin d'améliorer l'apprentissage organisationnel et développer une meilleure mesure de la contribution de la Banque au développement afin d'améliorer son efficacité sur le terrain* »[617]. Nous avons retracé la biographie de Robert Mc Namara et avons noté qu'il a été le concepteur de la méthode d'analyse des politiques publiques *Planning, Programming and Budgeting System* (PPBS) au sein du ministère de la défense américain dans les années 1960. Comme nous l'avons vu dans le premier chapitre, c'est le PPBS qui a conduit au développement de l'évaluation dans les politiques publiques américaines. Nous comprenons alors mieux son engouement pour construire un système d'évaluation au sein de la Banque mondiale, dont le premier des contributeurs était les États-Unis.

[616] Mervin Weiner, *Allocution lors du séminaire des 30 ans du département d'évaluation de la Banque mondiale*, archive de la Banque mondiale.
[617] Entretien avec une évaluatrice retraitée de la Banque mondiale, 20 novembre 2013.

La décision de Robert Mc Namara a élevé l'évaluation d'un sujet occasionnel de recherche au statut d'une activité régulière faisant partie de la routine de l'organisation. Les évaluateurs de l'époque sont unanimes quant à l'importance qu'a eu le soutien du président Robert Mc Namara dans l'essor de la fonction d'évaluation. L'un d'entre eux raconte, par exemple, « Mc Namara, était un supporter inconditionnel de l'évaluation. Son soutien était très clair dans tout ce qu'il disait et dans tout ce qu'il faisait. Il entretenait un lien direct et régulier avec le nouveau département d'évaluation par des réunions informelles et formelles. Si le président avait été moins intéressé par l'évaluation et plus distant, l'histoire de l'évaluation à la Banque aujourd'hui aurait été très différente. Tout le crédit lui revient »[618].

Robert Mc Namara a lancé un travail de sensibilisation sur l'évaluation auprès du personnel de la Banque. Toutefois, selon les évaluateurs de l'époque, il a « toujours eu de la peine à persuader les agents que les évaluations se feraient en toute objectivité »[619]. Pour réduire le risque de conflits d'intérêts et améliorer l'indépendance des évaluations, Robert Mc Namara a créé *l'Operations Evaluation Department* (OED), en 1973, dont il a donné la direction à Mohammed Shoiab, un des vice-présidents de la Banque, exprimant un grand intérêt pour l'audit organisationnel. En 1975, Mervin Weiner est devenu le second directeur du département et a mis l'accent sur l'importance de diffuser une culture de l'auto-évaluation chez le personnel de la Banque[620]. Son intention était d'augmenter l'apprentissage au sein des départements opérationnels, ainsi que de créer une division des tâches plus efficiente en matière d'évaluation entre, d'une part ces départements, qui seraient chargés d'élaborer les rapports de fin de projets et, d'autre part l'OED qui serait chargé de leur ratification et de la réalisation d'évaluations de plus grande envergure. Il a donc expérimenté une forme de décentralisation de la fonction d'évaluation qui deviendra quelques décennies plus tard une bonne pratique internationale. Cet accent mis, par Mervin Weiner, sur l'apprentissage et l'autoévaluation pourrait en partie s'expliquer par sa carrière passée[621].

Dans un premier temps, l'OED rendait directement des comptes au directeur budgétaire de la Banque, puis, au vice-président administratif. En moins de cinq ans, grâce à l'appui de Robert Mc Namara, le département de l'OED a été placé sous la responsabilité d'un directeur général, ce qui donnait à l'évaluation à un statut hiérarchique plus élevé.

---

[618] Susan Stout, Allocution lors du séminaire de la Banque mondiale célébrant les 30 ans de son département d'évaluation, archive de la Banque mondiale.
[619] *Ibid*.
[620] Archives de la Banque mondiale : *The World Bank / IFC Archives*, Transcript of the Interview with Mervin Weiner, 1986
[621] *Ibid*.

Au milieu des années 1980, la fin du mandat de Robert Mc Namara et le départ à la retraite du directeur de l'OED Mervyn Weiner ont marqué une baisse de l'intérêt pour l'évaluation au sein de la Banque mondiale. Un ancien évaluateur témoigne : « l'évaluation recevait beaucoup moins de soutien qu'au temps de Mc Namara qui jouait un rôle actif pour assurer que les évaluations soient valorisées et utilisées »[622]. Se sont alors succédés de nombreux directeurs de l'OED, témoignant de l'instabilité de cette fonction au sein de l'organisation. Le département souffrait d'un fort *turnover*, car la fonction d'évaluation était jugée peu attractive et était peu valorisée par la direction de la Banque. Graham Donaldson relate qu'« en 1984, le personnel de l'OED était mis à part. Il n'y avait pas d'esprit d'équipe. Beaucoup avaient été transférés contre leur gré à l'OED. Ils étaient offensés par le fait d'être placés dans le bâtiment M, un bâtiment reculé de l'organisation. L'OED était dans une situation difficile. Certains membres du département manquaient de compétences et la qualité des rapports n'était pas très bonne, voire parfois médiocre »[623]. Pour plusieurs anciens évaluateurs, c'est l'arrivée d'Yves Rovani, en tant que directeur général, qui a permis une redynamisation de l'OED. Dans des archives de la Banque mondiale, ce dernier déclare avoir été profondément admiratif du travail de Robert Mc Namara et très motivé pour poursuivre l'œuvre de ce dernier[624]. Une évaluatrice de la Banque mondiale rapporte : « *l'arrivée d'Yves Rovani, en 1986, a fondamentalement changé les choses. Le bureau de l'ancien directeur général était au septième étage et je pense ne jamais l'avoir vu à notre étage au sixième. Yves Rovani a vite installé le sien au sixième étage. Il profita de la réorganisation de la Banque, en 1987, comme d'une opportunité en or pour renouveler le personnel du département (...) Yves a permis de replacer le département d'évaluation au centre de la Banque en incitant les évaluateurs à travailler avec plus de rigueur et à être plus en contact avec les départements opérationnels. Il entretenait également des relations étroites avec le Joint Audit Committee, ce qui permit d'obtenir un meilleur soutien, du budget et du personnel* »[625].

Ce n'est toutefois qu'au début des années 1990 que les évaluateurs de la Banque mondiale situent le « renouveau » de l'évaluation au sein de l'organisation. L'évènement marquant a été la publication par l'OED d'un rapport de grande ampleur concluant à la baisse des performances des programmes mis en place par la Banque, à la suite duquel les dirigeants de la Banque ont remis en cause leurs certitudes sur les bonnes pratiques de l'aide au développement. Les

[622] World Bank, Oral History Program, *transcript of the interview with Graham Donaldson*, 2012.
[623] Graham Donaldson, *Allocution lors du séminaire organisé par la Banque mondiale pour les 30 ans de son département d'évaluation*, archives de la Banque mondiale.
[624] World Bank, Oral History Program, *transcript of the interview with Yves Rovani*, 2003.
[625] Entretien avec une évaluatrice retraitée de la Banque mondiale, 20 novembre 2013.

évaluateurs de la Banque racontent à ce sujet qu'ils étaient conscients depuis une décennie de la baisse de la qualité des interventions de la Banque sans arriver à avoir une influence sur la prise de décision[626]. Le retentissement de ce rapport est dû à ce sentiment latent au sein de la Banque que l'efficacité des programmes d'aide était en berne et faisait suite à plusieurs rapports faisant état de ce manque d'efficacité. Alors que les acteurs de la Banque identifient un « renouveau » brutal, il s'agit en réalité d'un changement incrémental qui s'est réalisé par l'accumulation de constats sur les lacunes de la Banque.

De plus, pendant de nombreuses années, l'inaction de la Banque mondiale face au constat de son inefficacité avait été critiquée par le monde académique et les ONG internationales travaillant dans le secteur de l'aide au développement[627]. Dans un ouvrage sur la gestion des projets d'aide au développement, Steven Brechin fait une analyse critique de la culture de l'apprentissage de la Banque mondiale dans les années 1980 : « la capacité des projets de développement à rencontrer les besoins des populations n'a pas été pour l'heure un souci majeur de la Banque tant que la survie des programmes de prêt n'était pas mise en danger. La Banque mondiale a géré les critiques de la communauté internationale en répondant par des changements structurels qui ont pourtant très peu affecté la performance de la Banque. La création d'un département d'évaluation n'a eu pratiquement aucun impact sur les opérations quotidiennes. C'est grâce à sa puissance symbolique et à sa réputation que la Banque mondiale a continué de croître malgré les impacts économiques, sociaux et environnementaux négatifs de beaucoup de ses projets »[628]. Cela rejoint l'analyse développée par Michel Crozier, selon qui « l'acteur attendra pour effectuer un changement qu'une dysfonction soit devenue assez grave pour menacer la survie même de l'organisation. Les décisions s'appliqueront alors même aux secteurs qui n'étaient pas sérieusement affectés par les dysfonctions »[629].

C'est alors que James Wolfensohn, président de la Banque mondiale de 1995 à 2005, a revalorisé la fonction de l'évaluation dans le but de donner à la Banque l'image d'une organisation qui apprend des leçons du passé. Il est présenté comme le premier président, depuis Robert Mc Namara, à donner une importance centrale à l'évaluation et à l'apprentissage : « il n'y avait pas eu de changement organisationnel si profond de la Banque depuis la présidence de Mc Namara. Dès le début de son mandat, Wolfensohn a travaillé à faire de la Banque une institution plus orientée vers l'apprentissage et la connaissance, plus concentrée sur la qualité,

[626] World Bank, 2002, *op. cit.*
[627] Voir par exemple, Brechin, Steven, 1997, *op. cit.* pp. 270-275. ou Easterly, William, PFUTZE, Tobias, 2008, *art. cit.*
[628] BRECHIN, Steven, 1997, *op. cit.*
[629] Crozier, Michel, 1963, *op. cit.* pp. 239-241.

plus réactive face aux attentes de ses tutelles financières, des pays emprunteurs, de son personnel et de ses partenaires du développement dans le secteur privé et la société civile »[630]. James Wolfenshon, avocat puis banquier avant de diriger la Banque mondiale, a fait de la lutte contre la corruption et de la transparence deux piliers de son mandat.

L'arrivée de ce nouveau président s'est conjuguée à la nomination de Robert Picciotto à la tête de l'OED, qui est resté dans l'histoire pour sa contribution à la transformation du département dans les années 1990 et pour sa revalorisation de l'évaluation au sein de la Banque. Dans les documents d'archive de la Banque mondiale, Robert Picciotto déclare que son intérêt pour l'évaluation a été inspiré par la pensée d'Albert Hirschman, avec qui il a pu s'entretenir dans les années 1970, rencontre qui l'a profondément marqué. Il fut également inspiré par le travail de Robert Mc Namara. Nommé au département d'évaluation, il réussit à convaincre James Wolfensohn du rôle fondamental de l'évaluation dans l'amélioration de la gestion de la connaissance et de l'apprentissage au sein de la Banque. Un évaluateur témoigne : « Bob Picciotto savait que pour que le travail de l'OED soit pertinent, il fallait qu'il s'aligne sur les nouvelles orientations de la Banque. Il avait conscience que l'OED ne pouvait plus seulement prendre les projets comme unité de mesure de l'efficacité des interventions de la Banque. Bob Picciotto a rapidement élaboré une nouvelle stratégie pour l'OED, afin que la direction de la Banque l'approuve et l'intègre dans la réforme plus générale de la Banque. Le personnel a critiqué le caractère non-participatif de la démarche.

Cependant, selon une évaluatrice de la Banque mondiale, « *une telle approche était nécessaire, car il fallait positionner rapidement le département de l'évaluation dans la nouvelle organisation de « banque du savoir conçue par James Wolfensohn, afin qu'il y joue un rôle important* »[631]. Robert Picciotto a réformé les pratiques du département d'évaluation en décidant, par exemple, que la détermination du programme annuel des évaluations se ferait en collaboration avec les directeurs opérationnels. Il a également œuvré à faciliter l'utilisation des évaluations (avec des évaluations alignées sur le temps de la décision politique, des rapports diffusés sous des formats concis pour un accès plus facile aux principales recommandations, etc.). Ce rôle central de Robert Picciotto est reconnu aussi bien au niveau du département de l'évaluation que par les autres membres de la Banque : « Bob Picciotto a personnellement contribué à la littérature de l'évaluation et a joué un rôle clé dans le renforcement des méthodes

[630] Patrick G Grasso, *Allocution lors du séminaire de la Banque mondiale célébrant les 30 ans de son département d'évaluation*, archives de la Banque mondiale.
[631] Entretien avec une évaluatrice de la Banque mondiale, 24 octobre 2013.

d'évaluation de la Banque », ou encore : « Bob Picciotto a été le DGO qui a permis une troisième renaissance de l'OED »[632].

Cette analyse montre que la diffusion d'une culture de l'évaluation à la Banque mondiale est avant tout le résultat de la volonté de certains acteurs qui ont développé, de par leurs expériences passées ou après avoir été inspirés par des prédécesseurs à la Banque mondiale, un intérêt particulier pour l'apprentissage et l'évaluation dans leur approche de la bonne gouvernance des programmes d'aide au développement. Les périodes où ce volontarisme n'est plus incarné, la fonction d'évaluation s'affaiblit. Nous remarquons également que le symbolisme du positionnement physique du département au sein de l'organisation est très fort, dans l'imaginaire collectif, pour matérialiser la stature d'une fonction. Quelques acteurs clés ont encouragé et accompagné la mise en place de structures facilitant la diffusion d'une culture de l'évaluation et de l'apprentissage dans l'organisation, les ont promues auprès de la Direction et des départements opérationnels, ont défendu le besoin en budget et en ressources humaines du département de l'évaluation et ont ainsi permis sa montée en importance au sein de la Banque.

En dépeignant l'influence des acteurs dans l'institutionnalisation de la fonction d'évaluation au sein de la Banque mondiale, nous espérons avoir montré l'importance d'incarner les processus de diffusion de bonnes pratiques au sein des organisations. Étudier ces processus à partir des acteurs permet de mieux comprendre les dynamiques qui les ont animés. Nous défendons la thèse selon laquelle si des facteurs organisationnels facilitent ou freinent le développement d'une expertise en évaluation au sein d'une organisation, ce sont avant tout les interactions entre acteurs qui permettent sa légitimation et son utilisation. Ce n'est qu'une fois la fonction d'évaluation profondément enracinée dans la culture organisationnelle que le portage des acteurs s'affaiblir sans menacer la pérennisation de cette expertise, mais ce n'est pas encore le cas aujourd'hui car l'évaluation a encore une légitimité fragile[633].

- **Des évaluateurs en quête d'identité**

[632] Archives du séminaire organisé par la Banque mondiale en 2000 pour les 30 ans de son département d'évaluation.
[633] PICCIOTTIO, Robert, "Evaluation and the Disciplines in Development", In VAESSEN, Jos, LEUUW, Frans. L, *Perspectives on Policy Evaluation and the Social Sciences,* Transaction Publishers, New Brunswick, 2010, 435p.

Selon le CAD-OCDE, « le rôle des départements d'évaluation est maintenant clairement défini dans la plupart des agences d'aide. Le problème de la position institutionnelle de l'évaluation dans les agences a été largement résolu par les nouvelles politiques ou documents cadres formulés par les unités d'évaluation »[634]. Nous remettons partiellement en question cette affirmation. Si donner une légitimité institutionnelle à l'évaluation permet l'intégration de l'évaluation dans la routine bureaucratique des organisations, cela ne suffit pas à garantir que le rôle d'acteurs de changement des évaluateurs sera légitimé par les autres acteurs des organisations.

Si l'affirmation du CAD-OCDE peut être vraie dans le cas de la Banque mondiale, dans la majorité des organismes d'aide, les évaluateurs ont encore du mal à se positionner, à promouvoir une vision claire de leur rôle et à concilier les différentes tâches qui leur sont assignées. Les témoignages d'évaluateurs que nous avons recueillis montrent une mise en avant de la fonction d'apprentissage de l'évaluation[635]. Une évaluatrice de la Banque mondiale déclare : « *l'évaluation c'est l'apprentissage. L'apprentissage à propos de la Banque et ce qui fait son efficacité* »[636]. Ce discours pourrait être expliqué par la culture organisationnelle de la Banque qui favorise la diffusion interne de la connaissance. Toutefois, nous retrouvons ces propos dans les témoignages d'évaluateurs d'agences bilatérales, où la culture de l'apprentissage est moins institutionnalisée comme à la JICA : « *le rôle premier des évaluateurs est de tirer des leçons des programmes d'aide* »[637]. Le fait que la mise en œuvre des évaluations soit externalisée permet aux évaluateurs internes de prendre de la distance avec la fonction de redevabilité de l'évaluation qui est remplie par les évaluateurs externes. Les premiers ne sont alors plus directement les « censeurs » de leurs collègues, ce qui rendrait plus difficiles leurs relations. Ils peuvent alors se présenter comme participant à l'apprentissage interne, une fonction plus « valorisante » auprès de leurs collègues. Toutefois, ces derniers ne font souvent pas vraiment la distinction entre les évaluateurs externes qui seraient les « censeurs » et les évaluateurs internes qui diffusent des leçons apprises. Ces deux types d'acteurs sont vus comme participant à la même tâche de contrôle.

Nous avons souligné auparavant que les unités d'évaluation sont souvent « isolées » des départements opérationnels et rendent compte aux conseils administratifs des agences d'aide, ou leur équivalent, et non directement aux instances de décision par soucis d'indépendance. Ce

[634] DAC-OECD, 2010, *op. cit.* p. 19.
[635] Ce sentiment ressort régulièrement dans les témoignages d'évaluateurs séniors dans les organismes d'aide au développement qui supportent mal l'externalisation et la décentralisation de la fonction d'évaluation.
[636] Entretien avec une évaluatrice de la Banque mondiale, 5 juin 2013.
[637] Entretien avec un évaluateur de la JICA, 8 mai 2012.

positionnement montre que les organisations attendent davantage des évaluations qu'elles donnent des résultats rigoureux, quantifiés, simples et faciles à communiquer dans une logique de redevabilité, plutôt que des analyses complexes permettant d'améliorer les pratiques de l'aide et de comprendre les causes des réussites et échecs des programmes. Les évaluateurs sont conscients de la signification de ce positionnement. Une évaluatrice de l'UNESCO déclare ainsi : « *le rôle de l'évaluateur dépend beaucoup de la structure organisationnelle dans laquelle il est placé. Ici, l'évaluation est située avec l'audit et l'investigation et cela joue sur l'image de l'évaluation et sa fonction. L'évaluation est trop déconnectée des unités opérationnelles. Les agents de l'UNESCO et les évaluateurs ont une volonté de tirer des leçons des évaluations, de comprendre ce qui fonctionne et ne fonctionne pas. Cependant, il y a une forte pression pour rendre compte des résultats. L'évaluation est plus proche d'une activité de management que de recherche* »[638].

Au-delà des contraintes organisationnelles qui ne facilitent pas la circulation des leçons tirées des évaluations, nous constatons des incohérences entre la volonté des évaluateurs de contribuer à l'apprentissage des acteurs de l'organisation et la responsabilité qu'ils estiment assumer dans ce processus d'apprentissage. Les agences d'aide et les évaluateurs ont aujourd'hui conscience que la simple diffusion des rapports, quels que soient leur qualité et leur intérêt, ne suffit pas à entraîner un apprentissage et une appropriation des recommandations par les différents acteurs concernés par l'évaluation. En conséquence, des mécanismes de suivi des recommandations des évaluations ont été mis en place. Toutefois, notre observation participante à l'AFD nous a permis de constater qu'un flou persiste sur qui porte la responsabilité de l'opérationnalisation de ces mécanismes. Le débat est vif entre les évaluateurs sur la question de savoir s'il est ou non de leur ressort d'assurer le suivi des évaluations. Pour certains, la diffusion des leçons apprises des évaluations serait de la responsabilité des directeurs des départements opérationnels[639]. Un membre de la division d'évaluation explique la réticence des évaluateurs en ces termes : « *la mise en place de la diffusion des résultats des évaluations à l'intérieur et à l'extérieur de l'agence se confronte au problème de la fatigue de l'évaluateur. Une évaluation complexe dure en moyenne 18 mois, et à la fin, les évaluateurs sont fatigués, voire écœurés. Ils sont donc assez réticents à se charger de l'effort de diffusion. Cette difficulté*

[638] Entretien avec une évaluatrice de l'UNESCO, 6 avril 2012.
[639] Observation faite lors de réunions d'équipe de l'unité d'évaluation de l'AFD auxquelles nous avons pu participer lors de notre observation participante à l'AFD en 2012 et 2013.

pourrait être dépassée si l'AFD mettait en place des incitations pour motiver les évaluateurs à mener cet exercice, mais elles sont inexistantes aujourd'hui »[640].

Devant la faible volonté de certains évaluateurs, et d'autres acteurs des agences d'aide, à assurer le suivi des évaluations, les mécanismes de suivi restent très procéduraux et bureaucratiques, sans qu'il y ait suffisamment d'interactions entre les acteurs. Une étude de la SIDA sur la qualité de ses évaluations est révélatrice à ce sujet : « même si des mécanismes de suivi des évaluations de performance sont disponibles, leur usage est souvent hésitant et partiel (…). Leur utilité est remise en cause. De plus, les départements opérationnels et stratégiques montrent un intérêt considérable à apprendre des *success stories*, mais il y a une réticence générale à analyser les échecs en profondeur. L'identification des bonnes pratiques est souvent faite sous l'influence des institutions influentes (Banque mondiale, CAD-OCDE) et non à partir de l'analyse de la propre expérience des acteurs »[641]. Cette citation montre la tendance des agences d'aide bilatérales à intégrer les bonnes pratiques promues par les organisations multilatérales ou par leurs pairs, sans les adapter à leur propre culture organisationnelle.

Les évaluateurs expriment toujours des difficultés à valoriser leur production intellectuelle et à la transmettre aux autres départements des agences d'aide. De nos entretiens, il ressort que les évaluateurs partagent le sentiment que leurs organisations ne leur donnent pas un pouvoir de persuasion et de négociation suffisant pour que l'évaluation influence la prise de décision. Par exemple, un ancien directeur du département de l'évaluation à l'AFD déclare : « *l'évaluateur n'a pas d'autorité sur la décision. Les évaluations qui mènent à des changements radicaux sont des exceptions* »[642]. Un autre témoignage d'un évaluateur de la JICA est révélateur : « *la capitalisation de connaissances, le recours aux résultats des évaluations lors de la planification de nouveaux projets et l'influence des évaluations sur le travail des opérationnels et des dirigeants sont faibles. Les évaluateurs n'ont pas le pouvoir d'assurer que les évaluations soient prises en compte* »[643]. Nous voyons ainsi que, malgré leur positionnement organisationnel avantageux (haut placés dans l'organigramme et indépendants), les évaluateurs se sentent frustrés par le manque de pouvoir de persuasion et de légitimité que leur donne l'organisation. Michel Crozier et Erhard Friedgerb expliquent comment une organisation investit certains acteurs « d'une autorité légitime sur d'autres, c'est-à-dire de pouvoirs

[640] Discussion informelle lors de notre observation participante à l'AFD.
[641] SIDA, 2008, *op. cit.*
[642] *Ibid*.
[643] Entretien avec un évaluateur de la JICA, 28 novembre 2014.

particuliers de sanction et de récompense. En d'autres termes, elle met aux mains de certains des atouts propres qui peuvent donner à ceux-ci un poids plus grand dans les négociations »[644]. Dans le cas des évaluateurs, c'est bien l'absence d'instruments de communication, d'incitation et de sanction qui les empêche d'exercer une influence sur les autres acteurs de l'organisation. Nous avons constaté, en effet, qu'hormis peut-être dans les organismes d'aide anglo-américains, l'évaluation n'est pas une fonction « de pouvoir »[645].

Les dirigeants des agences d'aide attendent également que les évaluateurs jouent un rôle de représentation de leur organisation à l'extérieur lorsqu'ils dialoguent avec les représentants d'autres agences lors d'évaluations ou qu'ils participent à des colloques internationaux. Cette fonction de l'évaluateur, qui n'était pas formulée quelques décennies auparavant, est maintenant régulièrement citée par les évaluateurs lorsqu'on les interroge sur leur rôle. Par exemple, une évaluatrice de l'AFD, déclare que : « *les évaluateurs sont la courroie de transmission entre l'extérieur et l'intérieur de l'agence* »[646]. Un autre rôle que l'on retrouve dans les discours des jeunes évaluateurs et peu dans celui des évaluateurs « senior » et retraité est de devoir contribuer au renforcement du partenariat entre les agences d'aide qu'ils représentent et les autres donateurs (dans le cadre des évaluations conjointes) et des bénéficiaires (dans le cadre des évaluations participatives). Une évaluatrice du DFID déclare : « *les évaluateurs s'attachent à faciliter le dialogue entre le bénéficiaire et le donateur et entre les donateurs* »[647]. Enfin, les évaluateurs des agences d'aide sont de plus en plus amenés à participer à des conférences internationales sur l'évaluation, où ils font valoir l'expertise et les bonnes pratiques de l'organisation qu'ils représentent[648].

Nous estimons que ces nouveaux rôles donnés aux évaluateurs, ou qu'ils se voient incarner, ont pris de l'ampleur à mesure que la fonction d'évaluation a été externalisée et décentralisée. Ne mettant plus en œuvre les évaluations, les évaluateurs se sont (vus) attribués d'autres rôles de substitution. La fonction d'évaluateur dans les agences d'aide évolue donc au fil du temps.

Certains évaluateurs notent que « *parallèlement à l'externalisation, l'activité d'évaluation s'est intellectualisée et conceptualisée. Les directeurs des unités d'évaluation et*

[644] CROZIER, Michel, FRIEDBERG, Erhard, 1977, *op. cit.* p. 146.
[645] Constats fait lors de notre observation participante à l'AFD et lors de dialogues informels avec des évaluateurs d'agences d'aide européenne lors de séminaires sur l'évaluation organisés par la société française d'évaluation, voir évènements observés en annexe 9.
[646] Entretien avec une évaluatrice d l'AFD, 8 octobre 2013.
[647] Entretien avec une évaluatrice du DFID, entretien Skype, 12 juillet 2013.
[648] Nous avons notamment pu organiser de telles conférences lors de nos différentes missions auprès de l'AFD.

les dirigeants des agences manquent de pragmatisme. Il faut que l'évaluation reste opérationnelle »[649]. En effet, les évaluateurs se sont vus de plus en plus engagés dans un travail de réflexion sur les normes, l'éthique, la méthodologie et les règles de procédures des évaluations. Les unités d'évaluation se sont mises à produire des « politiques » d'évaluation, des systèmes de notation de la qualité des évaluations, ou encore des travaux de recherche sur la méthodologie des évaluations d'impact.

Face à un rôle qui évolue rapidement et qui prend de multiples facettes (gestion, communication, recherche, etc.), les évaluateurs déclarent régulièrement qu'il leur manque des compétences pour remplir toutes les tâches qui leur sont demandées. Une étude du CAD-OCDE, réalisée en 2010, révèle ainsi que 60 % des unités d'évaluation des agences d'aide membres ont la perception qu'en plus de souffrir d'un manque de personnel, leurs compétences techniques ne sont pas suffisantes ou adaptées[650]. L'évaluation étant une fonction encore peu professionnalisée dans le secteur de l'aide au développement, il n'existe pas de standards établis concernant les compétences requises des évaluateurs, qui augmentent à mesure que la fonction d'évaluation prend de l'ampleur.

Cette faible professionnalisation favorise le fait que les évaluateurs soient principalement recrutés en interne et ont précédemment travaillé dans les départements opérationnels, soit sur une zone géographique, soit en tant qu'expert thématique (économie, genre, environnement, etc.)[651]. Les seuls évaluateurs qui ont une expérience significative en évaluation avant de rejoindre l'unité sont les quelques recrutés externes. Les évaluateurs se forment donc de manière *ad hoc* auprès de leurs collaborateurs et dans les formations organisées par les organisations multilatérales, notamment la Banque mondiale. En l'absence de formations académiques diplômantes spécialisées en évaluation des politiques publiques dans de nombreux pays (tous exceptés les pays anglo-américaines)[652], c'est également le cas pour les nouveaux arrivants dans les cabinets d'évaluation des politiques publiques[653].

Peu de travaux académiques analysent les compétences des évaluateurs, même dans les pays anglo-américains où l'évaluation est davantage professionnalisée. La principale étude sur

[649] Entretien avec un ancien directeur du département d'évaluation de l'AFD, 17 septembre 2013.
[650] DAC-OECD, 2010, *op. cit.*
[651] Ce constat a pu se faire à partir de l'étude des curriculum vitae, disponibles en ligne, d'une quarantaine d'évaluateurs d'une dizaine d'organismes d'aide. En annexe 10, nous présentons les profils de certains des évaluateurs, internes et externes aux agences d'aide, que nous avons interrogés.
[652] Il en existe, depuis peu, en France, principalement le Master de Sciences po Lyon en évaluation des politiques publiques créé par le directeur du cabinet d'évaluation Euréval basé à Lyon. Ces formations restent toutefois très rares.
[653] MATYJASIK, Nicolas, *op. cit.*

ce thème a été menée par Laurie Stevahn *et al*. en 2005[654]. Elle énumère une soixantaine de compétences nécessaires à un évaluateur. Ces compétences regroupent des connaissances générales en évaluation (théorie, méthodologie, standards internationaux, etc.), la capacité à mener des évaluations (conception des questions évaluatives, mise en place d'une méthodologie qualitative ou quantitative adaptée, etc.), une capacité à analyser le contexte social, politique et organisationnel de l'évaluation, des compétences en matière de management (communication, négociation, encadrement d'une équipe, etc.), ou encore la volonté de contribuer à l'amélioration de la professionnalisation de l'évaluation. Toutefois, ces quelques études[655] ne traitent pas spécifiquement des évaluateurs dans le secteur de l'aide au développement et restent très théoriques. Pour pallier ce manque, nous avons tenté d'observer les compétences attendues des évaluateurs dans les organismes d'aide et de les mettre en rapport avec leurs capacités réelles[656].

Nous distinguons plusieurs caractéristiques attendues des évaluateurs dans les processus de recrutement externe[657] : des compétences techniques sectorielles (dans la santé, l'eau, l'agronomie, etc.), une connaissance de la zone géographique du programme évalué, une connaissance de l'institution commanditaire (permettant une meilleure compréhension de la culture et des procédures de l'institution), une connaissance des théories et des méthodes de l'évaluation (approche quantitative et qualitative) et des compétences de recherche, notamment pour les évaluations d'impact. La pondération entre ces différentes compétences dépend surtout de l'offre. Dans des pays comme la France, où les formations en évaluation sont rares, les connaissances sectorielles ou géographiques en matière de développement sont priorisées. Ce ratio sera plus équilibré, voire inversé, dans les pays où l'évaluation est plus professionnalisée comme aux États-Unis ou au Canada. Les capacités attendues des évaluateurs recrutés en interne sont quelque peu différentes : nous retrouvons l'exigence d'une bonne connaissance de l'organisation, mais les connaissances préalables en matière d'évaluation ne sont pas un prérequis, étant donné qu'il est très rare que les personnels aient été en situation de se former à

[654] STEVAHN, Laurie, *et al*. « Establishing Essential Competencies for Program Evaluators », *Evaluation*, vol. 26, n°1, mars 2005.

[655] RUSS-EFT, Darlène, *et al*, *Evaluators Competencies, Standards for the Practice of Evaluation in Organizations*, John Wiley and Son, 2008, 238p; DEWEY, Jennifer. D, "Evaluators' Competencies, What's Taught Versus What Sought's", *Amercian Journal of Evaluation*, vol. 29, n°3, septembre 2008, pp. 268-287; SCRIVEN, Michael, "Types of Evaluators and Type of Evaluations", *Amercian Journal of Evaluation*, vol. 17, n°2, 1996, pp. 61-71.

[656] Afin d'explorer quelles compétences étaient attendues des évaluateurs dans les agences d'aide nous avons consultés les « fiches de postes » diffusés par des agences pour le recrutement de nouveaux évaluateurs. Nous avons également pu consulter la direction des ressources humaines de l'AFD. En ce qui concerne les compétences « réelles » des évaluateurs, nous nous sommes basé sur leurs curriculums vitae ainsi que sur des entretiens lors desquels nous leur avons demandé quelles capacités ils pensaient avoir et devoir améliorer. Notre observation participante à l'AFD nous a également permis de voir comment les évaluateurs travaillaient et quelles compétences ils mobilisaient au quotidien.

[657] Ce cas est relativement rare, dans la mesure où les postes sont généralement ouverts en premier lieu au recrutement interne, puis dans un second temps seulement, ouverts au recrutement externe si aucune candidature n'est présentée en interne.

l'évaluation. Les compétences requises des évaluateurs sont donc très variées : capacités sectorielles, spécialisation sur un type de financement (micro-crédit, aide budgétaire), compétences en management, en économie, etc. Nous avons réalisé une étude des parcours académiques et professionnels des évaluateurs internes et externes aux agences d'aide et avons découverts qu'ils sont très diversifiés. Les évaluateurs ont des profils davantage spécialisé sur le développement social, l'environnement, les relations internationales que d'économistes ou de financiers, ce qui les distingue des auditeurs[658].

Cette indétermination des compétences des évaluateurs donne l'impression aux autres acteurs du développement que l'évaluation n'est pas réellement une expertise ou une profession et que tout le monde peut s'improviser évaluateur[659]. Cette impression est renforcée par le fait qu'un bon nombre d'évaluateurs internes n'ont jamais mené une évaluation ou été formés avant d'entrer au département d'évaluation. L'expertise des évaluateurs est donc vue avec scepticisme.

Ce scepticisme remonte aux années 1980. À la Banque mondiale, les décideurs et les opérationnels estimaient que l'évaluation appartenait « à l'histoire » et était « hors du temps » de la prise de décision opérationnelle et politique. Ainsi, les opérationnels et les dirigeants se référaient peu au travail des évaluateurs, dont ils ne voyaient pas la valeur ajoutée : « la réputation du département de l'évaluation était d'être une sorte de maison de retraite. C'était là où nous allions à la fin de la journée pour consulter les dossiers de projets achevés. Ce n'était pas le service le plus en pointe. Le département de l'évaluation était aussi considéré comme conservateur, ne voulant pas changer sa manière de fonctionner »[660]. Un opérationnel retraité d'un autre organisme d'aide raconte également : « *je n'avais pas connaissance de ce qu'était l'évaluation. Je ne pensais vraiment pas que les retours du département de l'évaluation avaient un intérêt pour nous, les opérationnels* »[661]. Ainsi, pendant plusieurs décennies, l'évaluation fut perçue comme peu utile d'un point de vue opérationnel et décisionnel. Elle était en dehors du temps du politique. Si aujourd'hui, grâce aux réformes organisationnelles et aux efforts de professionnalisation et d'institutionnalisation de la fonction d'évaluation, l'utilité de l'expertise des évaluateurs est davantage reconnue, elle reste encore fragile.

[658] Voir annexe 10.
[659] Cette opinion nous a été régulièrement confiée durant nos enquêtes au sein d'organismes français d'aide au développement.
[660] Paroles recueillies dans les archives de la banque mondiale dans des retranscriptions d'interviews d'agents de la Banque avant leur départ de l'organisation dans le cadre du programme *World Bank Oral History.*
[661] Discussion informelle avec un acteur retraité de l'AFD au siège de l'AFD le 12 février 2013.

Nous constatons également que les discours du personnel des organismes d'aide sur les évaluateurs est plus positif lors des périodes où cette fonction est mise en valeur par la direction. Ainsi, dans les archives de la Banque mondiale, nous retrouvons des opinions plus favorables des décideurs et opérationnels sous les présidences de Robert Mc Namara et de James Wolfensohn.

Au-delà du soutien de la direction, le statut donné aux évaluateurs par les autres acteurs dépend de la durée depuis laquelle la fonction d'évaluation est formellement institutionnalisée. Si à la Banque mondiale et dans les banques de développement en général, le rôle central de l'évaluation est reconnu[662], cela est moins vrai dans les agences où cette institutionnalisation est externalisée ou plus récente, même si elles pratiquent l'évaluation de manière intensive. Au sein de la JICA, par exemple, il ressort de témoignages que nous avons pu recueillir[663] que le personnel de l'agence ne donne pas une importance centrale à l'évaluation en termes d'apprentissage et de prise de décision. De manière générale, comme nous l'avons déjà brièvement évoqué ci-dessus, nos observations et de nos entretiens dans les agences d'aide bilatérales montrent que l'évaluation n'est pas encore perçue comme une « expertise » à part entière. Le sentiment généralisé est que l'évaluation est une compétence de « généraliste », que les évaluations n'apportent que des conclusions qui sont « déjà connues » des opérationnels ou des dirigeants. Un sous-directeur financier du ministère des affaires étrangères parlait, par exemple, des évaluateurs comme de « *ceux qui sont placés là car ils ne sont pas capables d'être efficaces dans l'action* »[664]. Une ancienne évaluatrice d'un organisme français d'aide au développement déclarait avoir quitté son poste en grande partie car l'expertise des évaluateurs y était trop peu reconnue, ce qui rendait son travail peu épanouissant[665].

Dans les organismes d'aide, les évaluateurs se confrontent à des experts, hautement qualifiés, qui ont une croyance forte en leurs bonnes pratiques[666], ce qui les rend hermétiques à la remise en question de leur manière de faire. Dans de tels contextes, les évaluateurs ont des difficultés à faire valoir leur expertise lorsque celle-ci remet en cause les préconceptions des acteurs évalués. Cette remise en cause de l'expertise des évaluateurs est moins vraie en ce qui concerne les évaluateurs spécialisés dans les méthodes expérimentales. En effet, ces dernières sont reconnues comme demandant une expertise poussée en économétrie et un savoir-faire

[662] Entretien avec une évaluatrice retraitée de la Banque mondiale, 20 novembre 2013 ; Entretien avec un évaluateur de la Banque mondiale, 21 mai 2014.
[663] Entretien avec un évaluateur de la JICA, 8 mai 2012.
[664] Entretien avec un ancien agent du DFID travaillant désormais au ministère des affaires étrangères, 19 juin 2014.
[665] La personne interrogée souhaite rester anonyme sur ses propos précis.
[666] PRITCHETT, Lant, 2002, *art cit.*

difficile à maîtriser. Si des personnels internes n'ayant pas de compétences spécifiques en évaluation peuvent être recrutés à des postes d'évaluateurs « lambda », les agences d'aide recherchent des compétences poussées en méthodologie expérimentale d'évaluation en ce qui concerne les postes dédiés aux études d'impact.

Cette réflexion rejoint notre constat préalable sur la faiblesse de la culture de l'apprentissage dans les agences d'aide. Lant Pritchett, ancien économiste à la Banque mondiale, soutient, en outre, que même dans une perspective de redevabilité, réaliser une évaluation n'est pas la stratégie la plus optimale pour les acteurs des agences d'aide. Selon lui, l'utilisation des évaluations à des fins de redevabilité est « le choix de la faiblesse », c'est-à-dire le choix fait lorsque les programmes d'aide souffrent d'une forte remise en cause politique. Tant que la survie du programme n'est pas remise en cause par l'opinion publique ou par un changement d'orientation politique, le responsable du programme aura tout intérêt à avoir recours à la persuasion et au *storytelling*[667], plutôt qu'à avoir recours à des évaluations coûteuses, qui portent le risque de révéler des dysfonctionnements et donc d'ébranler un programme jusqu'alors perçu comme un succès[668]. Selon cette logique, la plupart des évaluations seraient donc menées pour répondre à la bonne pratique internationale de systématisation de l'évaluation, mais elles ne seraient pas utilisées par les opérationnels et dirigeants. Nous trouvons cette analyse excessive. Nous avons pu constater en effet que les évaluations sont aujourd'hui un élément indispensable de redevabilité et de communication avec l'opinion publique et que les méthodes quantitatives, qui s'éloignent du *storytelling*, sont privilégiées.

Il n'en demeure pas moins que les acteurs des agences d'aide développent des stratégies pour conserver le pouvoir d'influence de leur expertise sectorielle ou géographique par rapport aux évaluateurs. Des jeux de pouvoir dans la réalisation de travaux d'évaluation ce qui en fait un enjeu politique au sein des agences d'aide. L'approche la plus courante que les praticiens et les experts sectoriels ou géographiques mettent en œuvre est de critiquer la rigueur de la

[667] Le concept de *storytelling* provient du secteur de la communication et signifie l'art de raconter des histoires, des contes, afin de persuader les membres d'une organisation, l'opinion publique ou une clientèle ou bien pour créer un sentiment d'appartenance au sein d'un groupe. Un des premiers à utiliser ce concept dans le cadre des organisations internationales est Stephen Denning, ancien manager à la Banque mondiale. Dans son ouvrage *The Springboard How Storytelling Ignites Action in Knowledge-Era Organizations*, il raconte comment il a convaincu le personnel de l'organisation des bienfaits du "knowledge management" en usant du *storytelling* : « à cette époque dans l'organisation, le « knowledge management » était une idée étrange et incompréhensible pour le personnel de la Banque mondiale. J'ai essayé de les convaincre de faire de la Banque une « banque de savoir » par tous les moyens classiques de communication sans succès. En désespoir de cause, j'ai eu recours au *storytelling* et j'ai compris le pouvoir des histoires qui parlent souvent mieux que les chiffres, les théories ou les graphiques » (pp. 2-3).

[668] PRITCHETT, Lant, "It pays to be ignorant: A simple political economy of rigorous program evaluation", The journal of policy reform, 2002, pp. 251-269.

méthodologie des évaluations pour en affaiblir la fiabilité des conclusions. Par exemple, un agent de la direction de la Banque mondiale admet que « de manière souvent injustifiée, les agents de la Banque remettent en cause la méthodologie de l'évaluation pour disqualifier une expertise qui critique leurs pratiques habituelles et leurs conceptions du développement »[669]. Au sein des Nations unies, les responsables de projets et opérationnels trouvent tellement d'excuses pour ne pas être soumis à l'évaluation que les évaluateurs ont diffusé une liste humoristique de ces excuses[670]. Cette résistance à l'évaluation est en premier lieu à l'impression des agents des organismes d'être submergés par des exercices d'évaluation, d'être constamment soumis à une voire plusieurs évaluations[671].

Par ailleurs, les évaluations émettent un jugement de valeur argumenté sur les méthodes d'action et les résultats du personnel des agences d'aide. Si l'objectif d'apprentissage est mal perçu ou mis à l'arrière-plan par rapport à celui de redevabilité, alors les acteurs évalués peuvent se sentir « accusés » par l'évaluation. De plus, les évaluateurs soulèvent des débats sur le bien-fondé de certaines bonnes pratiques et politiques d'aide et participent ainsi à l'exacerbation de luttes de pouvoir entre les tenants des différentes approches. Nous développons la thèse selon laquelle la mise en place de règles et de procédures pour rendre les évaluations « neutres et objectives » ne permet pas de faire disparaître les rapports de force qui préexistent au sein des organisations. En effet, les relations entre les évaluateurs et les différents acteurs des agences d'aide restent fondées sur la négociation et le compromis, malgré l'existence de bonnes pratiques sur le contenu et la méthodologie des évaluations. En effet, pour que l'évaluation soit utilisée, il faut que ces recommandations soient acceptables par toutes les parties prenantes. Un compromis doit donc être trouvé entre ces parties prenantes et entre l'évaluateur et les acteurs du programme évalué pour que le rapport et ses recommandations restent « objectifs » mais qu'ils présentent des solutions acceptables et applicables. L'externalisation des évaluations dans la plupart des agences bilatérales était censée permettre de distancier les évaluateurs internes par rapport à ces jeux de pouvoir. Toutefois, dans la mesure où ce sont eux qui portent ces travaux en interne, ils s'y retrouvent inévitablement impliqués.

Étant donné que les évaluations sont principalement commanditées par les donateurs à des fins de redevabilité et de légitimation, ces derniers se retrouvent face à un dilemme entre le besoin d'évaluations « objectives et rigoureuses » et l'envie d'influencer le travail de

[669] Prem Garg (directeur assurance qualité), Joanne Salop (vice-présidente), Jan Piercy (directeur exécutif des États Unis), *Allocution lors du séminaire de la Banque mondiale célébrant les 30 ans de son département d'évaluation*, archives de la Banque mondiale.
[670] Cette liste d'excuses est présentée en annexe 5.
[671] Opinion recueillie auprès des agences de terrain de l'AFD au Laos, en Côte d'Ivoire et au Sénégal.

l'évaluateur pour renforcer la légitimité de l'organisation. Les chercheurs qui ont travaillé sur l'évaluation de l'aide au développement estiment que c'est la deuxième partie de ce dilemme qui est favorisée, ce qui entraînerait un adoucissement par les évaluateurs des conclusions des évaluations, afin de satisfaire les donateurs qu'ils présument préférer des évaluations positives[672]. L'évaluation serait donc biaisée et instrumentalisée pour satisfaire les décideurs. Comme évoqué ci-dessus, cet avis est partagé par certains évaluateurs seniors qui remettent en cause l'indépendance des évaluateurs externes. Toutefois, notre recherche a conduit à une analyse quelque peu différente. En effet, la plupart des évaluateurs internes que nous avons interrogés estiment qu'il leur serait difficile de conserver une posture indépendante de par leur position « interne » aux organismes d'aide et qu'ils ne se sentent pas à l'aise à l'idée de juger le travail de leurs collègues[673]. L'externalisation serait donc une bonne chose. Les évaluateurs externes défendent, quant à eux, l'indépendance de leur analyse : « *l'externalisation permet un gain important d'indépendance. Nos rapports établissent des conclusions et des recommandations auxquelles ne se risqueraient pas des évaluateurs internes. Dans de nombreuses évaluations, nous faisons état de difficultés de compréhension et de coordination entre les organismes de développement français que ces derniers refusent souvent d'entendre. Nos conclusions entraînent des tensions qui seraient difficiles à assumer pour des évaluateurs internes ayant vocation à travailler dans des départements opérationnels ou stratégiques de ces organismes dans le futur* »[674].

Notre recherche nous a conduit à développer l'idée que ce n'est pas l'influence des donateurs sur les évaluateurs externes qui fait que les évaluations semblent souvent prôner un *statu quo* politique, conduisant à pérenniser les rapports de force entre acteurs, mais le fait que les rapports soient « lissés » en interne avant d'être publiés. Selon un directeur de division à l'AFD, « *étant donné que les évaluations sont utilisées à des fins de redevabilité, les rapports sont « arrondis » et « lissés » avant d'être diffusés à la hiérarchie et vers l'extérieur. C'est pour cela que l'impression générale est que les rapports d'évaluation n'apprennent rien de nouveau* »[675]. C'est donc la pression pesant sur les donateurs pour qu'ils fournissent des preuves

[672] MICHAELOWA, Katharina, BORRMANN, Axel, "Evaluation Bias and Incentive Structures in Bi-and Multilateral agencies", *Review of Development economics*, vol. 10, n°6, 2006, pp. 313-329.
[673] Ces propos sont issus de conversations informelles avec les évaluateurs de l'AFD lors de notre observation participante en 2012 et 2013.
[674] Entretien avec le directeur d'un cabinet d'évaluation des politiques publiques basé en France, 9 décembre 2014.
[675] Entretien avec un directeur de division de l'AFD, 5 décembre 2013.

de bons résultats à leurs tutelles financières qui tend à faire que les rapports d'évaluation sont biaisés positivement lors de leur diffusion[676].

Le pouvoir des responsables de programmes sur l'évaluateur externe s'exprime surtout dans la mise en œuvre de l'évaluation, puisque sa coopération à l'évaluation est un élément fondamental à sa bonne réalisation. Le responsable de programme peut faciliter le travail de l'évaluateur en lui fournissant des contacts et des informations pertinents sur le programme évalué. Il a aussi le pouvoir de sélectionner ces contacts et ces informations, afin de cacher des problèmes potentiels dans la gestion du projet ou du programme. L'évaluateur doit donc gagner la confiance des acteurs de l'aide pour accéder à l'ensemble des données nécessaires à une analyse compréhensive des enjeux des projets et programmes évalués. Selon les économistes allemands Katharina Michaelowa et Axel Borrmann, « il est évident que l'objectivité d'une évaluation ne dépend pas seulement du caractère indépendant de l'évaluateur par rapport à l'agence. Les évaluateurs œuvrent sous une pression budgétaire et temporelle très stricte et apprécient beaucoup de pouvoir travailler dans un climat convivial avec les responsables de programmes. Ces derniers peuvent faciliter le travail des évaluateurs en lui arrangeant des entretiens et des visites de terrain, en présélectionnant la documentation pertinente et en ayant préparé les réponses aux questions que l'évaluateur pourrait avoir. Grâce à ce soutien, l'évaluateur gagne du temps, mais il risque d'avoir une perspective moins objective sur les résultats du programme. Le responsable du programme, quant à lui, gagne la possibilité de réduire la rigueur et la qualité de son travail sans encourir le risque que cette attitude se reflète par des résultats négatifs des évaluations »[677]. Nous estimons que si une perte d'objectivité involontaire est toujours possible, l'adoption volontaire d'un tel comportement est relativement rare, car cela est risqué aussi bien pour l'évaluateur que pour l'opérationnel qui peuvent être sanctionnés (mauvaise réputation, refus de paiement, limogeage) si le stratagème est révélé. De plus, les responsables de programmes auraient le même rôle de facilitateur, qui est indispensable, envers des évaluateurs internes.

Ainsi les tensions sont nombreuses entre évaluateurs externes, les évaluateurs internes et les décideurs et opérationnels des agences d'aide. En effet, l'expertise des évaluateurs est porteuse de changement politique et est, dès lors, soumise aux luttes de pouvoir internes aux agences d'aide. Pour les acteurs de l'aide, légitimer les évaluateurs et leur expertise revient donc à devoir agir en fonction de leurs recommandations. Dans des organisations où

[676] CLEMENTS, Paul, "Monitoring and Evaluation for Cost-Effectiveness in Development Management", *Journal of Multidisciplinary Evaluation,* 2005, pp. 11-38.
[677] MICHAELOVA, Katharina, BORRMANN, Axel, 2006, *art. cit.*

l'évaluation est davantage appréhendée comme un instrument de redevabilité que d'apprentissage, ce pouvoir de l'évaluation peut être vu comme un danger, car l'évaluation devient un instrument de contrôle des bonnes pratiques des décideurs et des opérateurs des programmes d'aide. Ces derniers vont alors tenter de remettre en cause la légitimité des évaluateurs. En conséquence, les évaluateurs vont devoir entreprendre un travail de légitimation qui passe par un plaidoyer démontrant l'utilité des évaluations.

C. Utilité et usages de l'évaluation

Nous avons vu que l'évaluateur et son expertise avaient des difficultés à trouver une légitimité dans les organismes d'aide au développement. Ce sentiment pourrait être dû à un décalage entre les attentes d'information de changement que portent les donateurs sur les évaluations et les effets que celles-ci sont véritablement capables d'apporter sur les pratiques et les politiques d'aide.

▪ Un usage souvent incrémental et peu perceptible

Michael Patton est reconnu comme étant le concepteur de la théorie de « l'évaluation fondée sur l'utilité » (*utility focused evaluation*). En 1978, il publia un ouvrage défendant l'idée selon laquelle c'est l'utilité d'une évaluation qui en fait la qualité[678]. Dans un entretien, il déclare : « une manière de distinguer les différentes approches de l'évaluation est de regarder ce qu'elles considèrent comme fondamental pour mesurer la réussite d'une évaluation. Pour moi, c'est que l'évaluation soit conçue et mise en œuvre de manière à ce qu'elle puisse contribuer significativement à l'amélioration des programmes et de la prise de décision »[679]. La pierre angulaire d'une évaluation réussie est, selon lui, « la détermination des utilisateurs directs de l'évaluation et le travail fait avec eux pour identifier et permettre les usages de l'évaluation qu'ils attendent »[680]. Les évaluations devraient donc être conçues en fonction des besoins des publics visés. Leur qualité dépendrait de l'utilité des résultats et des recommandations, ainsi que du caractère inclusif et participatif de leur méthodologie, plutôt que de leur rigueur scientifique et de l'objectivité des conclusions formulées. Les travaux de Patton sont repris par

[678] PATTON, Michael, Q, *Utilization focused evaluations*, Sage Publication, 1978, 688p.
[679] PATTON, Michael, Q, entretien pour l'agence canadienne de développement, consulté le 3 novembre 2013.
[680] *Ibid.*

les études qui ont suivi sur l'utilisation des évaluations[681], auxquelles les organismes d'aide au développement n'ont fait référence que très tardivement.

Il est frappant de constater que la question de l'utilité des évaluations n'apparaisse qu'au milieu des années 2000 comme une préoccupation pour les agences d'aide, alors que la fonction d'évaluation était souvent opérationnelle depuis trois décennies. Basil Cracknell explique cela par le fait que « les évaluateurs ont été si préoccupés par le développement de leurs programmes d'évaluation et par la résolution des problèmes de méthodes et de procédures, qu'ils ont laissé de côté l'étude de la manière dont ils pouvaient utiliser et transférer les connaissances issues des évaluations le plus efficacement possible »[682]. Nous remarquons effectivement que c'est au moment où les standards internationaux d'évaluation ont été intégrés aux activités routinières des agences d'aide que les évaluateurs se sont penchés sur la question de l'utilité des évaluations. Cette observation est un exemple de la perte de sens que peut entraîner la production internationale de bonnes pratiques, lorsque le seul objectif des acteurs devient de les reproduire le plus fidèlement possible, sans qu'il y ait de réflexion sur leur utilité. Nous constatons également que l'apparition de l'enjeu de l'utilisation des évaluations apparaît au moment où sont publiés des travaux de recherche condamnant le faible apprentissage dont font preuve les agences d'aide bilatérales[683]. Les départements d'évaluation des agences d'aide ont alors commencé à mener des études pour démontrer l'utilité des évaluations qu'ils menaient[684] en se référant aux typologies établies par les chercheurs en politiques publiques s'intéressant à l'usage de la recherche dans la prise de décision politique. Nous nous réfèrerons ici, à la typologie de Carol Weiss[685] qui offre, selon nous, une analyse claire des fonctions potentielles de l'évaluation. À partir de cette typologie, nous analyserons comment les évaluations sont réellement utilisées dans les agences d'aide.

Le premier type d'usage est « instrumental ». D'après ce modèle, l'évaluation est utilisée de manière directe par les décideurs et opérationnels des agences d'aide pour éclairer leur prise de décision ou pour améliorer la mise en œuvre des programmes. Alors qu'en ce qui concerne les travaux de recherche, cet usage direct est peu attendu par les dirigeants des agences d'aide,

---

681 Par exemple : ALKIN, Marvin. C, "Naturalistic Study of Evaluation Utilization", *New Directions for Program Evaluations: Utilization of Evaluative Information,* vol. 5, 1980, pp. 19-28; GUBA, Egon. G, LINCOLN, Yvonna. S. *Fourth Generation Evaluation,* SAGE Publications, 1989, 249p; WEISS, Carol, "Have we learned anything new about the use of evaluation", *American Journal of Evaluation,* vol. 19, n°1, 1998, pp. 121-133.

682 CRACKNELL, Basil, E, 2000, *op. cit.* p. 185.

683 Par exemple, les ouvrages et articles de William Easterly cités en bibliographie ou encore, CGDev, *When will we ever learn ? Improving lives through impact evaluation,* Report of the evaluation gap working group, mai 2006.

684 Par exemple : AFD, *Les évaluations sont-elles utiles ? Revue de littérature sur connaissances et décisions,* janvier 2009 ; SIDA, 2008, *op. cit.*

685 WEISS, Carol, 1983, *art. cit.*

nous retrouvons une attente forte de mise en opération directe des travaux d'évaluation[686]. Les évaluateurs ont, eux-aussi, longtemps mis cette utilité instrumentale en avant, jusqu'à ce que la capacité des évaluations à avoir un impact direct sur l'amélioration des politiques et des pratiques de l'aide soit largement remise en cause[687]. Il a été admis par la plupart des évaluateurs que « même les meilleures évaluations affectent seulement minimalement la prise de décision »[688] et que leur utilité est davantage incrémentale que directe. Toutefois, nous constatons un regain d'enthousiasme pour l'approche de l'utilisation directe dans les années 2000, avec l'engouement généralisé pour les évaluations expérimentales. En 2004, une étude de Rossi *et al* défendait l'idée que « les recherches récentes sur l'utilisation des évaluations remettent en cause la croyance selon laquelle les évaluations n'ont pas d'utilité directe. Il semble en réalité qu'un nombre significatif d'évaluations soit l'objet d'une utilisation directe »[689].

Selon nous, il est avéré que les évaluations peuvent avoir un usage instrumental au niveau opérationnel lorsque les parties prenantes du projet ont pris part au processus d'évaluation et que, grâce à cette participation, elles sont convaincues du bien-fondé des recommandations qu'elles ont contribuées à formuler. Nous défendons l'idée que ce n'est pas la lecture des rapports, mais l'expérience concrète de l'évaluation qui fait que les acteurs d'une organisation lui donnent un sens. Nous nous rattachons ici à la conception de l'utilisation théorisée par Patton : « les changements comportementaux, d'orientation ou organisationnels sont le résultat de l'apprentissage qui prend place durant le processus d'évaluation »[690]. Dire que l'évaluation ne prend sens que dans cet exercice personnel au quotidien, c'est aussi dire combien elle ne peut se comprendre que dans son rapport au politique. L'évaluation, n'est pas un instrument, une technique, c'est un espace de la pratique politique et un lieu d'énonciation du politique[691].

Toutefois, cette utilisation directe, par les opérationnels, souvent au niveau local, est peu visible, d'autant plus lorsqu'il n'y a pas de mécanismes de suivi des recommandations, ou que les résultats de ce suivi ne sont pas rendus publics, ce qui est le cas dans la grande majorité des agences d'aide. L'usage direct des évaluations au niveau des instances de direction est plus rare, car les décisions stratégiques à l'échelle d'un pays, d'une région ou globale sont influencées

[686] AFD, 2009, *op. cit.*
[687] *Cf.* Chapitres 1 et 2.
[688] ALKIN, Marvin, *Debates on Evaluation,* Sage Publication, 1990, 304p.
[689] ROSSI, Peter H, et al, *Evaluation: A systemic approach,* septième édition, SAGE Publications, 2004, 380p, p. 33.
[690] PATTON, Michael, 1979, *op. cit.* p. 155.
[691] Affirmation qui s'inspire de l'analyse de la bureaucratie néolibérale de Béatrice Hibou (*op. cit.* p. 188).

par davantage de contingences (économiques, politiques, historiques). Le poids d'une évaluation isolée est alors plus faible[692].

Or, c'est au niveau politique et stratégique que l'utilisation de l'évaluation est la plus visible, car la communication autour de la prise de décision est plus grande. Le manque d'usage instrumental à ce niveau contribue donc à la perception généralisée que les évaluations sont sous-utilisées dans le secteur de l'aide au développement.

La fonction de l'évaluation la plus mise en avant aujourd'hui par les évaluateurs est une utilisation incrémentale ou cumulative. Selon ce modèle, les évaluations ont une influence sur les orientations politiques et stratégiques des dirigeants, mais cette influence se réalise sur le long terme grâce à l'accumulation de connaissances issues d'évaluations multiples sur un même enjeu. L'accumulation de « preuves » va alerter le décideur sur l'importance de changer d'orientation politique. L'effectivité de cette utilisation incrémentale est plus difficile à étudier que l'utilisation directe, dans la mesure où elle est diffuse. Toutefois, nous estimons que la faiblesse des mécanismes d'apprentissage des agences d'aide, étudiée ci-dessus, constitue un blocage à ce type d'utilisation.

En effet, une utilisation incrémentale des évaluations demanderait qu'il existe une mémoire des travaux d'évaluation dans les agences d'aide, ainsi que des mécanismes incitant les responsables à se référer aux évaluations passées dans leur prise de décision. Lors de notre observation participante à l'AFD et grâce à des entretiens menés auprès d'évaluateurs, nous avons constaté que cette mémoire des évaluations réalisées est très faible. Les travaux d'évaluation sont exécutés de manière fragmentée et ne sont consultés que durant un court intervalle de temps après leur publication. De plus, les évaluations apparaissent aux yeux des responsables de programmes comme des travaux très contextualisés difficilement applicables à d'autres contextes. Les acteurs de l'aide ont tendance à considérer leur cadre d'action comme très spécifique et à douter fortement de la connaissance que pourrait leur apporter des évaluations sur des programmes similaires mais non-identiques[693]. Ainsi, les départements opérationnels sont souvent sceptiques quant à la capacité des exercices de synthèse d'évaluations portant sur un même instrument, une même zone géographique ou un même secteur, à leur apporter des enseignements utiles pour les programmes dont ils ont la charge. Un autre facteur affaiblissant l'usage incrémental des évaluations, relevé par un nombre

[692] Sauf lorsqu'il s'agit de projets ou programmes très politisés, qui mettent en jeu les relations diplomatiques entre deux pays dans le cas d'une coopération bilatérale.
[693] SIDA, 2008, *op. cit.*

significatif d'évaluateurs, serait la qualité très variable des évaluations[694]. Celle-ci rendrait difficile ces exercices de synthèse des évaluations. Cette mauvaise qualité, réelle ou prétendue, sert d'argument aux opérationnels pour remettre en cause la fiabilité des recommandations proposées par les évaluations

Toutefois, nous avons démontré que la solution à la trop faible utilisation des évaluations ne réside pas dans l'amélioration de la diffusion et du suivi des recommandations des rapports d'évaluation comme le suggère le CAD-OCDE[695], mais dans une implication plus grande des acteurs des agences d'aide aux processus d'évaluation. Cependant, cette participation active des acteurs des agences d'aide dans les évaluations ne peut avoir lieu que s'ils sont incités à le faire par leur organisation car il s'agit d'une activité chronophage. Selon une évaluatrice de la Banque mondiale, « *l'utilisation incrémentale des évaluations réside avant tout dans la culture de l'auto-évaluation qu'ont acquise les agents d'une organisation* »[696].

Le troisième type d'usage de l'évaluation est « persuasif » ou « tactique », c'est-à-dire que les acteurs impliqués dans le programme évalué vont commanditer et/ou utiliser les résultats de l'évaluation comme des arguments pour défendre le bien-fondé du programme, de la politique, ou encore de l'instrument évalué. Cet usage tactique de l'évaluation est assez tabou chez les évaluateurs, dans la mesure où il donne l'impression que l'évaluation est « instrumentalisée » à des fins politiques. Toutefois, dans un contexte où l'évaluation a pour principale fonction la légitimation des activités des organismes d'aide et la redevabilité, nous avons eu l'intuition que cet usage est plus important que ce que la littérature sur l'évaluation de l'aide au développement ne laisse penser. Cette importance se révèle notamment par le biais positif que plusieurs chercheurs notent dans la publication des évaluations et encore plus dans la citation des évaluations[697]. Dans nos entretiens, nous avons observé que les évaluateurs valorisent de plus en plus cet usage de l'évaluation, jusqu'alors considéré comme négatif[698]. Laurent Fontaine, directeur de la division de l'évaluation de l'AFD, déclare par exemple : « l'utilisation de type plaidoyer des évaluations est une bonne chose et ne doit pas être stigmatisée, car l'un des rôles de l'évaluation est aussi de défendre des programmes de manière

[694] Intervention de Laurent Fontaine, directeur du département d'évaluation de l'AFD à la Société française d'évaluation, novembre 2012 ; entretien avec une évaluatrice à la Banque mondiale, 25 octobre 2013 ; archives de la Banque mondiale.
[695] Solution souvent recommandée dans les revues par les pairs menée par le CAD-OCDE.
[696] Entretien avec une évaluatrice de la Banque mondiale, 25 octobre 2013.
[697] Par exemple : EASTERLY, William, 2008, *op. cit;* BRECHIN, Steven, 1997, *op. cit.* MICHAELOVA, Katharina, BORRMANN, Axel, 2006, *art. cit.*
[698] AFD, 2009, *op. cit.*

crédible et indépendante, chose que ne peuvent pas forcément faire les parties prenantes des projets »[699].

De plus, nous estimons que l'utilisation tactique et l'utilisation incrémentale ne s'opposent pas. En effet, la seconde se détermine pendant l'évaluation, alors que l'usage tactique a lieu *a posteriori*, lorsque les dévideurs vont présentent ou non les résultats au public et les utilisent ou non à des fins de légitimation. S'il y a un biais positif dans ce dernier type d'utilisation, il n'empêche pas qu'un apprentissage et une utilisation incrémentale aient eu lieu durant le processus évaluatif.

- **Le poids de la négociation dans l'utilisation des évaluations**

Quel que soit le type d'usage de l'évaluation, nous le considérons comme un processus durant lequel les acteurs de l'évaluation vont s'approprier les recommandations ou les leçons apprises, pour ensuite les utiliser dans un objectif de réforme (direct), d'apprentissage (incrémental) ou de justification (tactique). L'utilisation des évaluations est généralement présentée comme une simple mise en pratique des recommandations. Elle est pour nous un exercice plus complexe, où l'interprétation des acteurs, la culture organisationnelle, ou encore le contexte d'intervention vont avoir un poids important. Lorsque le sujet de l'utilisation des évaluations est abordé par la littérature sur le développement, il est peu fait mention des négociations entre acteurs qui précèdent l'adoption ou le rejet des mesures issues des recommandations. L'utilisation de l'évaluation est exclusivement étudiée à partir du nombre de rapports d'évaluation publiés et de l'existence de mécanismes de rétroaction pour assurer le suivi des résultats et des recommandations des évaluations. Dans notre recherche, nous avons voulu faire une analyse plus sociologique de la phase d'utilisation des rapports d'évaluation. En effet, durant cette phase, les rapports de force des différents groupes d'intérêts des agences d'aide vont se confronter pour orienter l'opportunité de réforme ouverte par l'évaluation (ou la « grappe » d'évaluations) en leur faveur.

Nous estimons que c'est à travers la multiplication des négociations et des échanges de points de vue formels et informels pendant les processus d'évaluation et lors de la revue des rapports que l'utilisation des évaluations pourra être renforcée. Selon Voivamirta Valovirta, « l'étape d'argumentation et de prise de décision est un processus interactif, durant laquelle les résultats (des évaluations) sont débattus, reconstruits et positionnés dans un contexte social

[699] Laurent Fontaine, allocution à la société française d'évaluation, novembre 2012.

précis. Elle combine la tentative de convaincre à un niveau cognitif, puis de persuader les autres à agir. Ces interactions peuvent aboutir à des décisions et à des actions, ou à une nouvelle compréhension partagée du problème. Elles peuvent également déboucher sur un renforcement ou un affaiblissement de la légitimité de l'institution »[700]. Cette citation illustre le fait que ce n'est pas la prise de décision qui est l'indicateur pertinent pour juger de l'utilité de l'évaluation, mais le dialogue et la réflexion collective qu'elle ouvre. Toutefois, cette utilisation ne peut être étudiée que de l'intérieur, par exemple par une observation participante de long terme au sein d'une organisation d'aide, permettant d'assister aux interactions qui ont lieu durant les processus d'évaluation. Notre observation participante à l'AFD et les dialogues que nous avons pu avoir avec des évaluateurs d'autres agences d'aide nous ont permis de constater que si les évaluations aboutissent rarement à des réformes de grande ampleur, elles sont souvent l'occasion de faire émerger de nouvelles problématiques et d'ouvrir un dialogue, qui n'aurait pas eu lieu autrement, entre les parties prenantes des programmes évalués.

Nous réfutons donc la croyance selon laquelle les évaluations ne sont pas utilisées par les agences d'aide et qu'elles seraient réduites à être une routine bureaucratique n'ayant pas d'influence sur les pratiques de l'aide au développement[701]. Alors que les évaluateurs, les donateurs et l'opinion publique attendent une utilisation concrétisée par des réformes, la réalité est que l'utilisation des évaluations est fragmentée au niveau des individus et peu visible, car apparaissant avant tout de manière cognitive par un changement progressif des préconceptions et des habitudes des acteurs. Les évaluations sont donc utilisées, bien qu'à des degrés différents selon l'intérêt des acteurs, leur qualité, l'importance des sujets traités dans l'agenda des agences d'aide, etc. Il manque surtout une synergie d'acteurs et des moyens pour que cette utilisation fragmentée devienne une utilisation « cumulative » et collective qui permette un apprentissage commun.

Suivant cette ligne de pensée, nous remettons en cause l'idée répandue selon laquelle une amélioration de la rigueur scientifique des évaluations, notamment par la mise en œuvre de méthodes expérimentales, entraînerait une hausse significative de leur utilisation. En effet, si la rigueur des processus évaluatifs et la qualité des évaluations sont primordiales à leur crédibilité, elles ne garantissent pas l'appropriation des recommandations et la compréhension de leur utilité par les parties prenantes des programmes évalués. À ce sujet, Michael Patton estime que la tâche la plus difficile de l'évaluateur est « d'aller des données aux recommandations. Lorsque

[700] VALOVIRTA, Voimavirta, « Evaluation Utilization as Argumentation », *Evaluation*, vol. 8, n°1, 2002, pp. 60-80, p. 65.
[701] BRECHIN, 1997, *op. cit.* ainsi que les analyse de William Easterly référencées en bibliographie.

vous collectez des données, vous regardez ce qui s'est passé. Mais lorsque vous écrivez des recommandations, vous êtes un futuriste. Les évaluations peuvent vous aider à faire des prévisions, mais les décisions futures ne sont pas juste une question de données. Un bon évaluateur peut talentueusement montrer les forces et les faiblesses d'un programme grâce à une collecte de données pertinentes et crédibles, mais cela ne garantit pas qu'il sache les transformer en recommandations utiles. (…) Je préfère impliquer les parties prenantes décisionnaires du programme évalué pour qu'elles génèrent leurs propres recommandations et que ces recommandations prennent en compte leurs valeurs, leur expérience et leurs ressources »[702]. Cette citation explique pourquoi le recours à des données quantitatives, de quelque qualité qu'elles soient, ne permet pas d'assurer qu'une évaluation sera utile. En outre, une approche purement expérimentale laisse peu de place aux outils d'analyse permettant des interactions entre les acteurs du projet (groupes de travail, focus group, entretiens collectifs, etc.), que nous avons identifiées comme étant un des vecteurs principaux d'utilisation des évaluations.

La littérature grise des organismes d'aide va dans le sens de cette idée. En 2008, la Banque mondiale a mené une étude qui émet un bilan mitigé sur l'utilisation des évaluations expérimentales : « on peut espérer que la connaissance produite soit plus facilement cumulative, mais la diversité des politiques et des contextes risque de limiter à quelques questions dispersées la création d'un réservoir de connaissances probant »[703]. Une autre étude menée par l'AFD conclut à une utilisation limitée des évaluations d'impact, que ce soit en termes de redevabilité, d'apprentissage, ou de justification des bonnes pratiques de l'agence[704]. Selon cette étude, la méthode expérimentale n'améliore pas l'utilisation directe des évaluations (son temps long de mise en œuvre rend difficile l'intégration des évaluations dans une « fenêtre d'opportunité politique ») et la rigueur des preuves apportées est souvent affaiblie par des défauts de procédures. En outre, leur utilisation incrémentale est remise en cause, car ces évaluations paraissent, encore plus que d'autres, seulement valables pour le projet spécifique qu'elles traitent, étant donné qu'elles impliquent une isolation de l'objet évalué. Leur principal effet s'exprime en termes d'utilisation tactique dans un contexte international où l'engouement politique pour ce type d'évaluation est fort. Les agences d'aide utilisent l'image de rigueur scientifique de ces évaluations pour convaincre de l'efficacité de leurs activités et persuader leurs tutelles financières de maintenir leurs financements.

[702] Michael Patton, entretien à l'ACDI, consulté le 3 novembre 2013.
[703] World Bank, 2008, *op. cit.*
[704] AFD, 2010, *op. cit.*

- **Conclusion de chapitre**

Ce chapitre nous a permis d'observer l'importance des acteurs dans le processus de diffusion d'une culture de l'évaluation, qui apparaît pourtant désincarné et très institutionnel dans la littérature sur l'évaluation de l'aide au développement. Les évaluateurs se trouvent au cœur de rapports de force qu'il s'agit de gérer et d'arbitrer. Ils sont eux même soumis à une forte remise en cause de leur expertise visant à les délégitimer et à contrôler leurs recommandations. L'établissement d'une culture de l'évaluation et la forme qu'elle prend dépendent avant tout des interactions entre acteurs qui se nouent autour des politiques et des pratiques d'évaluation. L'évaluation n'a pas de sens et d'utilité « en soi », ce sont les acteurs des organismes d'aide qui vont les déterminer. Ainsi, l'évaluation est un objet politique, dans la mesure où elle est déterminée par les rapports de force entre les évaluateurs et les autres acteurs de l'aide d'une part, et parmi ces derniers d'autre part. L'évaluation, lorsqu'elle apporte des changements directs ou incrémentaux dans les pratiques et les politiques d'aide, va à son tour modifier ces rapports de force ainsi que, peu à peu, les approches dominantes de l'aide. Elle peut aussi les légitimer et contribuer à leur pérennisation lorsqu'elle les valide par son jugement. Nous avons également vu que le statut des évaluateurs était encore fragile dans les organismes d'aide. L'expertise des évaluateurs est parfois remise en cause par les opérationnels qui peuvent y voir une menace. La spécificité de l'évaluation par rapport à l'audit et au contrôle n'est pas toujours bien comprise. À ce sujet, Robert Picciottio est réaliste : « l'évaluation se développe dans un environnement rude. En effet, sa place au soleil ne sera sécurisée que si elle reste pertinente face aux évolutions de l'aide au développement et démontre sa valeur ajoutée dans l'arène compétitive de l'apprentissage organisationnelle et de la redevabilité »[705]. Étudier l'évaluation à partir de ses acteurs nous a permis de mettre en avant le fait que la culture de l'évaluation dans les agences d'aide est encore un processus inachevé, malgré une pratique systématisée de l'évaluation, ce qui peut paraître paradoxal.

Dans ce chapitre, nous nous sommes concentrée sur le statut de l'évaluation du Nord au sein des organismes d'aide. Désormais, nous allons analyser l'émergence d'acteurs de l'évaluation au Sud. Nous allons étudier quels modes de transferts de compétences ont été mobilisés par les donateurs pour créer une expertise en évaluation au Sud. Nous allons nous demander si les évaluateurs du Sud ont su s'approprier les compétences importées par les

[705] Robert, Picciottio, intervention lors du séminaire organisé par la Banque mondiale pour les 30 ans de son département d'évaluation, archives de la Banque mondiale.

donateurs et si l'on peut parler aujourd'hui de système d'évaluation au Sud qui possèderaient des acteurs qualifiés et des structures opérationnelles.

Chapitre 5 : L'artifice du renforcement des capacités en évaluation au Sud

Nous aborderons, dans ce chapitre, les processus « verticaux » de diffusion d'une culture de l'évaluation des donateurs vers les pays bénéficiaires de l'aide au développement. Nous allons démontrer que les exigences des donateurs en matière de suivi-évaluation des programmes d'aide et les conditionnalités de bonne gouvernance qu'ils posent à leur allocation d'aide sont les principaux vecteurs qui incitent les États du Sud à développer des structures d'évaluation et à systématiser la pratique de l'évaluation. Pour aider les acteurs locaux (publics ou privés) à se doter de compétences en évaluation, les donateurs mettent en place des programmes de renforcement des capacités, et depuis peu, des évaluations partenariales. Nous allons analyser ces deux modes de transferts de compétences et leur efficacité pour réaliser une analyse critique de l'état des besoins et des compétences actuelles en évaluation dans les pays du Sud. Nous allons également étudier les relations qu'entretiennent les évaluateurs du Nord et du Sud et les caractéristiques sociologiques de ces derniers.

A. Des transferts de compétences inopérants

Quelques chercheurs se sont intéressés à l'analyse des causes de l'émergence de la pratique de l'évaluation dans un ou, de manière comparative, plusieurs pays. Dans leur *Atlas de l'évaluation*, Jean-Eric Furubo, Ray Rist et Rolf Sandhal[706] établissent une typologie des États selon les facteurs qui ont influencé leur décision de développer la pratique de l'évaluation dans leurs administrations publiques :

- les États ayant subi une forte pression interne[707] et une faible pression externe[708] : ce sont souvent les pays « pionniers » de l'évaluation et/ou qui se sont servis de l'évaluation pour justifier de grandes réformes économiques et sociales. Dans ce groupe d'États, la culture des politiques publiques était donc favorable à l'évaluation : États-Unis, Canada, Australie, Norvège ;

[706] FURUBO, Jean-Eric, RIST, Ray, SANDHAL, Rolf *op. cit.* pp. 21-22.
[707] Par pression interne, Furubo, Rist et Sandhal entendent une demande forte de la société civile.
[708] Par pression externe, ils entendent une incitation d'organisations supra-étatiques telles que l'Union européenne ou les organisations multilatérales. La pression externe peut aussi être plus indirecte et relever de la volonté des États de se conformer aux pratiques de leurs pairs.

- les États ayant été influencés par une forte pression interne et une forte pression externe : souvent des pays qui ont institutionnalisé l'évaluation dans les années 1970 et 1980, dans des périodes de crise économique et de restriction budgétaire et sous l'influence normative de leurs pairs : France, Royaume-Uni, Allemagne, Suède, Danemark ;
- Les États ayant été soumis à une pression extérieure forte sans qu'il y ait de demande nationale préalable : ce sont des pays qui ont établi tardivement une fonction d'évaluation sous la pression d'organisations multilatérales (Banque mondiale, Nations unies) et/ou régionales (Union européenne, banques régionales de développement). Leurs politiques publiques est souvent éloignée de la culture de l'évaluation, en raison d'une faible transparence dans l'administration, de problèmes de corruption ou d'une société civile faible : la Chine, l'Italie, l'Espagne, l'Irlande.

Notre recherche nous a amené à penser que les pays du Sud se classent, pour la plupart, dans la dernière catégorie avec, d'une part une pression externe forte des bailleurs de fonds pour qu'ils mettent en place des structures et une pratique d'évaluation et, d'autre part une pression interne faible due à des facteurs multiples (expertise indépendante en politiques publiques faible, corruption, faible culture de la transparence dans l'administration, société civile peu organisée, etc.). Les pays du Sud où l'impulsion de l'évaluation provient pour l'essentiel de facteurs internes sont rares et il s'agit souvent de pays émergents, comme le Brésil ou l'Afrique du Sud[709], ou dans une moindre mesure de pays dits à revenus intermédiaires, par exemple le Maroc[710]. Dans les pays où la culture de l'évaluation est presque complètement importée, sa diffusion sera plus lente et complexe que dans ceux connaissant une pression interne forte qui ont pu voir l'évaluation s'institutionnaliser en une dizaine d'années (États-Unis, Australie)[711].

Afin de modéliser les interactions qui sont en œuvre lors des processus de transferts de compétences et de standards normatifs d'évaluation entre les donateurs et leurs partenaires du Sud, nous reprenons la typologie d'Harold Wolman[712]. Cette typologie permet de souligner l'importance des acteurs dans les processus de transferts normatifs, ce qui rejoint notre approche

[709] World Bank, *Comment mettre en place des systèmes de suivi-évaluation pour améliorer les performances du secteur public*, 2007.
[710] Entretien avec un consultant en évaluation des politiques publiques au Maroc, 17 septembre 2014.
[711] FURUBO, Jean-Eric, RIST, Ray, SANDHAL, Rolf, 2002, *op. cit*,
[712] WOLMAN, Harold, 1992, *art. cit*,

sociologique de la diffusion d'une culture de l'évaluation des organismes donateurs vers les pays du Sud. Nous considérons ainsi que :

- les organismes multilatéraux et bilatéraux de développement sont les « inspirateurs » des « bonnes pratiques » en matière d'évaluation, c'est-à-dire les acteurs qui ont créé ces « bonnes pratiques » et qui ont l'ambition de les voir se diffuser à travers le monde ;
- les agents des organismes d'aide qui mènent des activités de renforcement des capacités dans les pays du Sud, ainsi que des officiels locaux directement impliqués dans ces activités, représentent les « opérateurs » de la diffusion des « bonnes pratiques d'évaluation » ;
- les dirigeants politiques locaux, ou encore les organisations de la société civile locales, qui font du plaidoyer pour une meilleure transparence et une gestion plus responsable des politiques publiques sont les « facilitateurs » qui créent un environnement favorable à la diffusion d'une culture de l'évaluation ;
- Enfin, les gouvernements et administrations des pays du Sud qui mettent en place des réformes en matière d'évaluation, les évaluateurs formés au Sud, et les populations qui bénéficient de l'évaluation, constituent les « receveurs » du transfert.

Les frontières entre ces catégories d'acteurs sont poreuses. Les mêmes acteurs peuvent remplir plusieurs fonctions. Les membres du gouvernement des pays du Sud peuvent, par exemple, être à la fois facilitateurs et receveurs, ou encore les évaluateurs du Sud sont à la fois opérateurs et receveurs du transfert.

- **L'historicité du renforcement des capacités en évaluation**

C'est essentiellement à partir des années 1990 que les organismes multilatéraux d'aide au développement se sont attachés à former des évaluateurs au Sud. Dans les décennies précédentes, des interventions ponctuelles de renforcement des capacités en évaluation prenaient place mais elles restaient anecdotiques. Ce sont la Banque mondiale et le PNUD qui s'emparèrent de cette thématique pour en faire un axe important de leurs programmes d'aide au développement.

La définition du « renforcement des capacités en évaluation » a été élaborée par le PNUD : « une intervention ou un ensemble d'interventions qui contribuent à l'établissement de capacités d'évaluation dans les structures administratives des pays bénéficiaires de l'aide »[713].

La volonté de généraliser les activités de renforcement des capacités en évaluation des pays du Sud était latente depuis le milieu des années 1970 dans plusieurs enceintes du développement, notamment au sein du CAD-OCDE, où elle était portée par les pays du nord de l'Europe et le Royaume-Uni sans entraîner d'enthousiasme de la part des autres pays membres[714]. Le nouvel intérêt pour cette thématique dans les années 1990 peut être en partie expliqué par le contexte politique de l'époque. En effet, la fin des années 1990 a été marquée par le retour d'approches de l'aide qui plaçaient l'État au centre des réformes à mener pour parvenir au développement. Ces approches étaient en rupture avec la période critiquée du Consensus de Washington et s'appuyaient sur de nombreuses études de chercheurs en économie du développement qui démontraient une corrélation positive entre un retour de l'État dans la régulation de l'économie et le développement[715]. Selon cette logique, les organismes d'aide au développement ont élaboré des programmes ayant pour objectif l'amélioration de la gouvernance des pays bénéficiaires de l'aide. C'est donc dans le cadre d'un large courant d'initiatives de renforcement des capacités institutionnelles (renforcement de l'État de droit, des systèmes exécutifs, judiciaires et législatifs, des capacités statistiques des administrations, etc.) que s'est inscrit le développement d'activités de renforcement des capacités en évaluation.

En outre, l'approche de la bonne gouvernance qu'avait adoptée la majorité des donateurs était largement influencée par la théorie du New Public Management présentant l'évaluation comme une bonne pratique de gestion des politiques publiques. Nos recherches ont montré qu'il n'y avait pas réellement de demande de renforcement des capacités en évaluation dans les pays du Sud. Cette demande provenait clairement des organismes d'aide au développement qui voyaient en l'évaluation une réforme positive pour les pays du Sud[716]. Les pays bénéficiaires de l'aide ont alors été fortement incités à reproduire cette demande et à introduire l'évaluation comme objet de réforme politique dans leurs plans nationaux de développement. Cette incitation passait notamment par le conditionnement de l'aide à des engagements des États à mettre en place des systèmes évaluatifs. Le développement de capacités en évaluation faisait partie du « package » de réformes institutionnelles que les donateurs proposaient aux pays

[713] KHAN, Adil, "Evaluation Capacity Building: An Overview of Current Status, Issues and Options", *Evaluation,* vol. 4, n°3, 1998, pp. 310-328.
[714] Entretien avec un membre du réseau d'évaluateurs du CAD-OCDE, 16 novembre 2012.
[715] Se référer à l'annexe 3.
[716] Entretien avec le directeur d'un cabinet d'évaluation basé en France, 20 novembre 2013.

bénéficiaires pour atteindre ce qu'ils considéraient comme étant un niveau minimal de bonne gouvernance permettant d'assurer l'efficacité des programmes d'aide.

Au sein de la Banque mondiale et du PNUD, la théorisation de méthodes de renforcement des capacités en évaluation et l'essor de cette activité furent favorisés par l'existence, au sein des départements d'évaluation, de promoteurs de cette thématique du développement. Quelques acteurs clés permirent que cette activité soit légitimée et qu'on lui attribue des moyens humains et financiers. Keith MacKay fut l'homme au cœur de la conception d'une approche spécifique à la Banque mondiale en matière de renforcement des capacités en évaluation[717]. Ainsi, la Banque mondiale put se positionner par rapport aux autres organismes de développement sur cette nouvelle thématique. Cet ancien fonctionnaire australien était connu pour avoir activement contribué au développement d'une culture de l'évaluation dans son pays dans un temps très court. Son objectif à la Banque mondiale était de modéliser, à partir de son expérience passée, une action de renforcement des capacités en évaluation qui soit reproductible dans tous les pays bénéficiaires des financements de la Banque mondiale. Son modèle, toujours utilisé par la Banque mondiale, est devenu un standard international. Il est constitué de deux phases[718] :

- l'élaboration d'un diagnostic pour identifier les fonctions de suivi-évaluation[719] qui existent déjà dans les pays, ainsi que leurs forces et leurs faiblesses du point de vue de la demande (soutien politique à la fonction évaluative, utilisation effective des informations du suivi-évaluation, etc.) et de l'offre (disponibilité d'une expertise en évaluation, qualité des rapports d'évaluations, existence de mécanismes de suivi des recommandations des évaluations, etc.). Le diagnostic doit, dans l'idéal, être réalisé de manière collaborative avec les pays partenaires, afin que les acteurs locaux soient sensibilisés à l'intérêt du suivi-évaluation ;
- la conception d'un plan d'action, dont la mesure principale consiste en la détermination d'un « champion » du suivi-évaluation, c'est-à-dire d'un acteur influent de la fonction publique (ministre, haut fonctionnaire) capable de devenir évaluateur, de lancer et soutenir le processus d'institutionnalisation du suivi-

[717] Entretien avec le directeur d'un cabinet d'évaluation basé en France, 20 novembre 2013.
[718] World Bank, *Evaluation Capacity Development, Lessons from National Experience,* ECD Working Papers serie n°1, OED, mai 1998.
[719] Lorsque l'on parle de renforcement des capacités en évaluation, cela recouvre en réalité également le suivi qui est vu comme une étape préalable primordial à la mise en œuvre d'évaluation. Les termes renforcement « en évaluation » et en « suivi-évaluation » sont donc utilisés indifféremment.

évaluation, de persuader ses collègues de la priorité qu'il constitue et du besoin qu'il y a d'allouer des ressources significatives à cette tâche[720].

Une place importante est donc donnée à l'acteur dans ce modèle théorique. Selon nous, l'idée de désigner un acteur « champion » de l'évaluation qui portera son institutionnalisation provint de l'expérience passée de Keith MacKay, mais également de l'histoire de la Banque mondiale, où l'évaluation fut portée par quelques « champions » et devint grâce à eux une fonction centrale de l'organisation[721].

Au PNUD, c'est essentiellement sous l'impulsion d'Adil Khan, spécialiste de la gouvernance socio-économique, que l'activité de renforcement des capacités en évaluation s'est développée. Il a théorisé une approche de cette activité qui se voulait divergente de celle de la Banque mondiale. Il estimait que la Banque mondiale cherchait à développer les capacités en évaluation de ses partenaires du Sud pour accroître indirectement son contrôle sur la gouvernance économique de ces derniers. Le renforcement des capacités en évaluation lui permettrait d'assurer « la qualité de son portfolio de prêts »[722]. L'objectif de la Banque mondiale aurait donc été d'améliorer sa propre redevabilité, plutôt que de générer un apprentissage en matière d'évaluation chez ses pays partenaires. Le PNUD déclarait vouloir proposer une approche plus centrée sur *l'empowerment* de la société civile, en orientant le développement de l'évaluation des politiques publiques vers une hausse de la transparence et de la redevabilité des gouvernements dans un objectif d'amélioration de la démocratie participative. Selon le PNUD, son appui à la diffusion d'une culture de l'évaluation dans les pays du Sud avait donc pour objectif premier de donner les moyens à la société civile de mieux contrôler et influencer la politique du gouvernement. Nous relevons, dans la pensée du PNUD, une influence des théories d'Armatya Sen sur le développement par les capabilités[723], économiste qui influença beaucoup la politique du PNUD dans les années 1990.

Pour renforcer les capacités en évaluation de ses pays partenaires, le PNUD a majoritairement mis en place des activités d'assistance technique portant sur le suivi-évaluation des projets qu'il conduisait. Ainsi, malgré les critiques du PNUD à l'égard de l'approche du renforcement des capacités en évaluation de la Banque mondiale, qui serait « centrée sur ses propres-intérêts », nous retrouvons également au PNUD une approche qui se focalisait sur

[720] World Bank, *Evaluation Capacity Development, Institutionalization of Monitoring and Evaluation Systems to Improve Public Sector Management*, ECD Working Paper Series, n°15, IEG, janvier 2006.
[721] *Cf.* Chapitre 4.
[722] KHAN, Adil, 1998, *art. cit.*
[723] NUSSBAUM, Martha, SEN, Amartya, *The quality of life*, Oxford University Press, 1983, 472p.

l'amélioration du suivi-évaluation des projets qu'il mettait en œuvre et qui répondait donc à ses besoins de redevabilité. Les projets ou les programmes de développement qui avaient pour objet principal le renforcement des capacités en évaluation étaient très rares. La Banque mondiale était, en réalité, le seul organisme à avoir développé de tels programmes. Dans les autres organisations multilatérales, le renforcement des capacités en évaluation n'était qu'une composante de programmes sectoriels plus larges.

À la fin des années 1990, la Banque mondiale a débuté la publication d'une série d'études sur le renforcement des capacités en évaluation, afin de fournir une expertise sur les bonnes et les mauvaises pratiques observées dans ce champ d'intervention après une décennie de mise en œuvre[724]. Elle a également élaboré les premiers guides sur la mise en œuvre de systèmes de suivi-évaluation à destination des pays bénéficiaires de l'aide et des autres donateurs[725]. Le PNUD, quant à lui, s'essaya à l'élaboration de monographies, afin d'établir un parangonnage des besoins et des capacités existantes en suivi-évaluation dans les États du Sud. La démarche était similaire à celle des « diagnostics-pays » de la Banque mondiale. Le PNUD organisait également des séminaires de formations régionaux, à la suite desquels il faisait des dons aux participants pour qu'ils continuent le travail de plaidoyer et d'élaboration de systèmes d'évaluation dans leurs pays.

Nous constatons donc que malgré les efforts affichés du PNUD pour présenter une approche différente de la Banque mondiale, ses interventions de renforcement des capacités étaient très similaires. Dans les deux cas, derrière une volonté, sûrement sincère, d'améliorer la gouvernance des pays du Sud, l'idée sous-jacente était de faciliter la réalisation d'évaluations de qualité des programmes d'aide. En effet, plus les structures et l'expertise d'évaluation locales sont solides, plus il est facile pour les donateurs de collecter les données quantitatives et qualitatives nécessaires aux évaluations. Ainsi, le renforcement des capacités se faisait toujours à partir de programmes de développement financés par ces organismes et servait leurs contraintes de redevabilité. La logique n'était donc pas purement « charitable », mais plutôt de type « gagnant-gagnant ». La seule distinction que nous identifions entre le PNUD et la Banque mondiale était la vision de l'évaluation qu'ils présentaient aux bénéficiaires de l'aide. La Banque mondiale mettait en avant les aspects de calcul « coût-bénéfice », de mesure de la

[724] Ces études sont consultables sur le site de l'Independent Evaluation Group de la Banque mondiale : http://ieg.worldbankgroup.org/ieg-search?f[0]=field_document_categories%3A334&f[1]=field_doc_sub_category%3A215 consulté le 21 novembre 2014.

[725] Ces « guides » sont consultables sur le site de l'Independent Evaluation Group de la Banque mondiale http://ieg.worldbankgroup.org/iegsearch?f[0]=field_document_categories%3A334&f[1]=field_doc_sub_category%3A339, consulté le 5 décembre 2014.

performance et de contrôle permis par l'évaluation, alors que le PNUD mettait davantage l'accent sur le potentiel de l'évaluation en matière de démocratisation des sociétés et d'émancipation des peuples.

Suivant l'exemple du PNUD et de la Banque mondiale, les pays membres du CAD-OCDE ont commencé à s'intéresser à la thématique du renforcement des capacités en évaluation. Les pays membres les plus engagés étaient logiquement les États anglo-américains et du nord de l'Europe (Cf. p. 239)[726]. Leurs activités de renforcement des capacités en évaluation prenaient surtout la forme de formations *via* l'organisation de séminaires régionaux sur l'évaluation. Les Banques régionales de développement ont aussi mis en place de l'assistance technique et des formations en matière d'évaluation[727].

La pratique du renforcement des capacités en évaluation a continué de se développer dans les organismes d'aide durant les années 2000. En effet, les bonnes pratiques établies dans le cadre de l'agenda sur l'efficacité de l'aide appellent à un renforcement des capacités en évaluation des pays du Sud. Par exemple, le principe « d'alignement sur les systèmes nationaux » implique que les donateurs recourent aux systèmes nationaux de suivi-évaluation, or, il faut pour cela que ces derniers soient suffisamment développés[728]. Le principe de redevabilité mutuelle[729] implique, quant à lui, que les pays bénéficiaires de l'aide évaluent la qualité des programmes d'aide des donateurs et qu'ils disposent donc de systèmes de suivi-évaluation opérationnels. Enfin, la « bonne pratique » de gestion axée sur les résultats[730] exige que les pays du Sud développent des systèmes de suivi-évaluation, afin d'être capables de mesurer la performance de leurs politiques publiques et de développer une culture du résultat. La Déclaration de Paris a ainsi placé le renforcement des capacités en suivi-évaluation en haut de l'agenda politique des agences d'aide au développement.

Alors qu'à la fin des années 1990, seuls quelques donateurs s'étaient investis dans le renforcement des capacités en évaluation, notamment les États-Unis, le Danemark, le Canada ou la Suède, ils sont plus d'une vingtaine aujourd'hui[731]. De plus, des ONG s'emparent de la question[732] et les Nations unies créent des réseaux d'évaluateurs spécifiquement dédiés à ce

[726] Entretien avec le directeur d'un cabinet d'évaluation des politiques publiques basé en France, 20 novembre 2013.
[727] RIEPER, Olaf, TOULEMONDE, Jacques, *Politics and Practices of Intergovernmental Evaluation,* New Brunswick, NJ: Transaction Publishers, 1997, 172p, pp. 130.
[728] CAD-OCDE, 2005, *op. cit.*
[729] *Ibid*.
[730] *Ibid*.
[731] DAC-OECD, *Fact-finding survey of evaluation capacity development (ECD) in partner countries,* octobre 2006.
[732] On peut par exemple citer *l'Evaluation Capacity Development Group*, une ONG États-uniennes qui appuie des États du Sud à renforcer leurs capacités en évaluation: http://www.ecdg.net/ , consulté le 24 novembre 2014.

sujet[733]. Toutefois, l'objectif de renforcement des capacités a souvent été rajouté de manière *ad hoc* à des programmes ou des réformes mis en œuvre à l'origine avec des finalités différentes[734]. Par exemple, la décentralisation des évaluations dans les agences d'aide est présentée par leurs dirigeants comme une mesure ayant été mise en place pour faciliter le renforcement des capacités en évaluation, puisqu'elle favorise l'embauche de consultants locaux et donc leur permet d'acquérir de l'expérience[735]. Mais, lorsque nous avons interrogé les évaluateurs sur les motivations de cette réforme et que nous avons consulté les documents préparatoires à son instauration, nous avons constaté que l'objectif d'origine était d'accroître le nombre d'évaluations réalisables et de réduire leurs coûts, et non la formation d'évaluateurs au Sud. Les agences d'aide raccrochent donc le renforcement des capacités à leurs activités pour se conformer aux bonnes pratiques des organisations multilatérales, sans pour autant développer de nouvelles actions spécifiquement dédiées à cet objectif. Les actions élaborées par les agences d'aide bilatérales dans un objectif premier de renforcement des capacités sont encore rares et ponctuelles.

Par ailleurs, nous constatons qu'au sein des Nations unies le mandat du renforcement des capacités en évaluation est passé du PNUD à l'UNICEF. En effet, le dirigeant du département d'évaluation de l'UNICEF, Marco Segone, s'est saisi de la thématique, alors qu'elle n'était plus portée par le PNUD depuis le départ d'Adil Khan. C'est après plusieurs années à avoir été responsable des évaluations au sein de bureaux régionaux et nationaux de l'UNICEF que Marco Segone a réalisé que le manque de capacités en évaluation dans les pays du Sud était un problème majeur pour leur développement. Selon lui, ce manque était un frein au renforcement des institutions et de la société civile. C'est ainsi qu'une fois nommé responsable de l'unité centrale d'évaluation de l'UNICEF à New-York, il s'est approprié cette thématique, profitant du vide sur la question au PNUD[736]. En plus, des actions traditionnelles de renforcement des capacités étudiées ci-dessus (assistance technique, formations, sensibilisation), Marco Segone a développé une stratégie orientée vers le soutien à la création de sociétés régionales et nationales d'évaluation et la constitution de réseaux internationaux d'évaluateurs, sur lesquels nous reviendrons dans le prochain chapitre.

[733] En 2009, l'UNEG a créé le *United Nations Evaluation Group for Asia and the Pacific: http://www.unicef.org/eapro/UNEDAP_Final.pdf*, consulté le 24 novembre 2014.
[734] Entretien avec le directeur d'un cabinet d'évaluation, consultant en évaluation auprès du CAD-OCDE, 20 novembre 2013 ; entretien avec la coresponsable du réseau d'évaluateurs IOCE, 22 janvier 2014.
[735] Dans les années 2000, les agences d'aide bilatérales suivent la tendance de décentraliser leurs activités d'évaluation de projets aux agences de terrain dans les pays bénéficiaires de l'aide.
[736] Entretien avec un évaluateur des Nations unies, proche collaborateur de Marco Segone, le 6 octobre 2014.

- **Une forme passive d'apprentissage inefficace**

Une enquête menée par le CAD-OCDE en 2006 montre que les activités de renforcement des capacités en évaluation des donateurs se sont longtemps focalisées sur la mise en place de formations et d'ateliers de travail, majoritairement à destination des hauts fonctionnaires des pays en développement[737]. En effet, le renforcement des capacités en évaluation a largement été considéré par les donateurs comme un enjeu de formation. Les formations internationales et régionales organisées ou financées par les banques régionales de développement, la Banque mondiale et les Nations unies se sont multipliées dans les années 2000, chaque organisme ayant mis en place sa propre formation internationale annuelle. La plus réputée et instituée est *l'International Program for Development Evaluation Training* (IPDET), co-organisée par la Banque mondiale et l'Université de Carleton au Canada. Ce programme fut créé en 2001 et vise à « fournir aux dirigeants et praticiens les outils de base pour évaluer les politiques, programmes et projets de développement aux niveaux national, régional et international »[738]. Le programme se déroule à Ottawa et est composé de deux semaines de cours et de deux semaines d'ateliers de travail animés par des chercheurs et évaluateurs du Nord et du Sud. La formation vise aussi bien un public du Nord (30 % des participants) qu'un public du Sud (70 % des participants).

Nous allons désormais démontrer que ces formations n'ont pas permis pas aux acteurs des pays du Sud de mettre en œuvre les compétences qu'ils ont acquises si une culture de l'évaluation ne s'est pas développée dans leurs pays d'origine. À cet effet, nous avons pu recueillir les témoignages de fonctionnaires africains ayant participé à des formations sur le suivi-évaluation organisées par des donateurs. Justine Odjoubé, directrice de l'Observatoire du changement social du Bénin, a participé à sept de ces formations depuis 2004 et nous a fait un retour sur ces expériences : « *comme principaux points forts, on peut noter les formations théoriques très bien ficelées. Par contre, en matière de formation pratique, peu de cas africains ou béninois ont été étudiés. À la suite de ces formations, nous devions nous exercer nous-mêmes à la pratique de l'évaluation, afin de combler les lacunes en ce qui concerne la formation pratique* »[739]. Un évaluateur sénégalais, déclare, quant à lui, que « *les formations internationales organisées par les organismes d'aide ne permettent pas de s'améliorer dans la pratique de l'évaluation. Toutefois, elles peuvent sensibiliser les participants à l'importance de l'évaluation pour améliorer la gouvernance dans leurs pays. C'est en faisant un Master dans une université française et un stage dans un cabinet d'évaluation en France que je me suis*

[737] DAC-OCDE (b), 2006, *op. cit.*
[738] Site Internet de l'IPDET/ http://www.ipdet.org/ consulté le 30 décembre 2013.
[739] Entretien avec la directrice de l'observatoire du changement social au Bénin, 2 décembre 2013.

formé à l'évaluation. La formation doit passer par la pratique et se fait sur le long terme »[740]. En effet, nous avons constaté que les évaluateurs du Sud sont nombreux à s'être formés et/ou à avoir fait des stages en évaluation des politiques publiques dans les pays occidentaux.

Certains évaluateurs du Nord sont, eux aussi, sceptiques quant à l'efficacité des formations internationales. Un évaluateur témoigne à ce sujet : « *je ne crois pas en l'efficacité de la formation pour renforcer les capacités, tout du moins pas dans les pays en développement. Après avoir reçu une formation, soit les individus partent rejoindre des organisations internationales ou régionales, soit ils restent dans leurs pays, mais alors ils ne pratiquent pas l'évaluation et leurs compétences se perdent. La formation n'est pas suffisante si elle est menée seule, elle ne mène alors à rien* »[741]. Ce type d'évènements est apprécié des donateurs, car il permet de rendre visible leur engagement sur la thématique du renforcement des capacités. Les partenaires du Sud s'y rendent avec plaisir car la formation théorique reste néanmoins instructive, qu'ils sont financés pour y assister et qu'ils y trouvent une opportunité de se créer un réseau professionnel international.

Nous avons pu participer en tant qu'intervenante aux formations à l'évaluation organisées par l'AFD. Nous avons constaté qu'elles restent cantonnées à des aspects très techniques de l'évaluation : réalisation des différentes phases d'une évaluation (élaboration de termes de références, lancement d'un appel d'offres, rédaction des questions évaluatives, des recommandations, etc.), respect des critères de qualité du CAD-OCDE, présentation des différents types d'évaluation avec un accent fort mis sur les évaluations d'impact, etc. Ces formations n'abordent ni les enjeux politiques qui accompagnent la diffusion d'une culture de l'évaluation (besoin de transparence dans la diffusion de l'information sur les politiques publiques, de programmes de lutte contre la corruption, d'un renforcement de la société civile), ni ceux qui relèvent de l'institutionnel (comment créer des liens entre la prise de décision et l'évaluation ? Comment assurer la prise en compte des recommandations des évaluations ?). Nous considérons que ces formations purement théoriques ne permettent pas de préparer les personnes formées à construire un contexte politique favorable à l'évaluation dans leurs pays.

Nous estimons donc que les organisations internationales font fausse route lorsqu'elles abordent le renforcement des capacités seulement par la voie de la formation. D'aucuns estiment que cette approche reflète la mentalité de « donneurs de leçons » des organismes

[740] Entretien avec le directeur d'un cabinet d'évaluation au Sénégal, 3 avril 2014.
[741] Entretien avec le directeur du cabinet d'évaluation des politiques publiques, consultant en évaluation auprès du CAD-OCDE, 20 novembre 2013.

d'aide[742]. Selon nous, ces formations sont un vecteur pour diffuser au Sud les bonnes pratiques que les organismes de développement ont produites en matière d'évaluation, ainsi que pour construire une approche harmonisée de l'évaluation au Nord et au Sud. Atteindre ces finalités serait plus difficile avec une autre forme de renforcement des capacités. Par exemple, si les organismes d'aide utilisaient une approche de type *learning by doing*, les acteurs formés ne seraient pas soumis aux discours sur les standards d'évaluation, ils pratiqueraient l'évaluation et construiraient leurs propres bonnes pratiques à partir de cette expérience et en fonction de leurs besoins et de leurs référentiels. Ainsi, il ne serait pas garanti que leur pratique de l'évaluation soit conforme à la bonne pratique internationale. Au contraire, dans les formations organisées par les donateurs, les acteurs du Sud sont abreuvés d'un discours théorique sur les standards qualitatifs de l'évaluation et sur ce que doit être une bonne évaluation. Lorsqu'ils mettront en œuvre des évaluations, il est alors probable qu'ils se conformeront automatiquement à cette approche de l'évaluation et qu'ils la véhiculeront dans leurs organismes d'origine. Cela est d'autant plus vraisemblable que les pays du Sud cherchent à réaliser des réformes politiques qui plaisent aux donateurs, afin d'attirer leurs financements.

Cette approche du renforcement des capacités centrée sur la formation révèle également que les donateurs s'intéressent davantage au renforcement de l'offre que de la demande en évaluation, considérant qu'il est très difficile d'influer sur cette dernière. Keith MacKay estime par exemple que « si la demande de suivi-évaluation est forte, il peut se révéler relativement simple d'améliorer l'offre pour y répondre, mais l'inverse n'est pas vrai »[743]. Nous pensons qu'il est possible d'agir sur la demande grâce à des activités de sensibilisation de l'administration sur les bénéfices de l'évaluation et en encourageant les gouvernements à réaliser des expérimentations d'évaluations, c'est-à-dire grâce au *learning by doing*. Nous estimons que pour diffuser une culture de l'évaluation, il est nécessaire d'agir simultanément sur l'offre, la demande et les structures organisationnelles. Ce point de vue est partagé par un certain nombre d'évaluateurs : « en tant que pratique innovante, l'évaluation croît par un processus interactif nécessitant une communauté professionnelle compétente du côté de l'offre et un groupe de consommateurs durable du côté de la demande. Il n'est pas réaliste de penser que l'émergence de l'évaluation doit être conditionnée à la préexistence d'une demande du secteur public. Au contraire, la demande doit être créée et progressivement développée au même rythme que l'offre, si ce n'est plus rapidement »[744]. Ainsi, les donateurs devraient donc

[742] Entretien avec le directeur d'un cabinet d'évaluation, consultant en évaluation auprès du CAD-OCDE, 20 novembre 2013.
[743] World Bank, 2006, *op. cit.*
[744] BOYLE, Richard, LEMAIRE, Donald, *Building Effective Evaluation Capacity: Lessons from Practice*, Comparative Policy Evaluation, Transaction Publisher, février 1999, 202p, p. 153.

agir sur la demande et sur l'offre. D'autant plus que la demande qui émerge spontanément, en dehors de tout contexte de réforme, est souvent superficielle et uniquement une réponse à une attente des donateurs.

Toutefois, en quoi consisterait une action sur la demande en évaluation ? Selon certains évaluateurs indépendants, il s'agirait de donner la capacité à la société civile d'exercer une pression sur le gouvernement pour que celui-ci engage des réformes allant vers une plus grande transparence des politiques publiques et un renforcement de la démocratie participative[745]. Nous remarquons donc que les évaluateurs externes aux organismes d'aide développent une approche « par le bas » du renforcement des capacités en évaluation, alors que les évaluateurs internes à ces organismes favorisent plutôt une approche intentionnelle « par le haut ».

Cette éviction du problème de la demande est notamment visible dans les « diagnostics pays » réalisés par la Banque mondiale qui, contrairement à ce qu'ils devraient être en théorie, portent exclusivement sur l'offre de suivi-évaluation[746]. De manière générale, la seule action qu'ont les organismes d'aide sur la demande en évaluation est l'organisation de séminaires internationaux ventant les bénéfices du suivi-évaluation aux gouvernements du Sud. Les comptes-rendus de ces séminaires dévoilent de véritables plaidoyers en faveur de l'évaluation : « l'évaluation peut être gratifiante pour un gouvernement qui la pratique de manière systématisée, car il sera vu par ses citoyens comme étant à l'écoute et protecteur »[747], ou encore : « les systèmes de suivi-évaluation sont au centre des dispositifs d'une saine gouvernance et sont une condition nécessaire à une gestion efficace des dépenses publiques en vue de réduire la pauvreté »[748]. L'efficacité de ces pratiques de plaidoyer est aujourd'hui fréquemment remise en cause par les chercheurs sur les transferts internationaux de politiques publiques[749].

Nous rejoignons ces critiques ; en effet, convaincre les hauts fonctionnaires nationaux des vertus d'un modèle n'entraîne pas automatiquement la mise en œuvre effective d'un transfert de ce modèle. La décision d'opérer le transfert et la manière dont cela va être fait dépendent, en premier lieu, des bénéfices que les « opérateurs » du transfert pensent pouvoir en tirer et de la capacité du contexte d'importation d'intégrer ce modèle[750]. Or, si les opérateurs ne

[745] Jacques Toulemonde, directeur du cabinet d'évaluation Euréval, consultant en évaluation auprès du CAD-OCDE, 20 novembre 2013 ; X, directeur d'un cabinet d'évaluation des politiques publiques, 13 janvier 2014.
[746] World Bank, 2007, *op. cit.*
[747] UNPD, *Evaluation capacity development in Asia, Selecting Proceedings from the International Conference Beijing,* octobre 1999.
[748] World Bank, *Influencing Change: Building Evaluation Capacity to Strengthen Governance*, 2011.
[749] DELPEUCH, Thierry, « L'analyse des transferts internationaux de politique publique : un état de l'art », *Questions de recherche*, n°27, CERI, Décembre 2008.
[750] *Ibid*.

sont pas déjà engagés dans un contexte réformateur visant à améliorer la redevabilité, la transparence et l'apprentissage au sein de la fonction publique de leur pays, ils se contenteront de mettre en place des structures d'évaluation (notamment car cela améliorera la réputation internationale des pays et attirera les financements internationaux des organismes d'aide) qui resteront des coquilles vides (car leur mise en fonction entraînerait une augmentation du contrôle public sur leurs activités, une charge de travail supplémentaire, etc.), même si leur opérationnalisation permettrait une meilleure efficacité de la gestion des politiques publiques. D'une part la notion de « meilleure efficacité » est toujours subjective, d'autre part ne nous situant pas dans un cadre théorique rationaliste, nous estimons que ce n'est pas parce qu'un acteur prend connaissance d'un modèle « efficace » qu'il va automatiquement vouloir l'adopter, notamment s'il entre en contradiction avec sa culture politique ou ses intérêts.

Ainsi, nous observons régulièrement dans les pays du Sud que les systèmes d'évaluation des politiques publiques semblent spécifiquement conçus pour répondre aux exigences des donateurs. Par exemple, aux Philippines, le *Project Monitoring Staff* de l'Autorité nationale d'économie et de développement est subdivisé par bureaux de donateurs (bureau de la Banque mondiale, bureau d'USAID, etc.) et non par secteurs de politiques publiques[751]. Un tel arrangement, même s'il est utile pour superviser les performances de l'aide, risque de dévier la priorité de l'évaluation des besoins du gouvernement, vers les besoins des donateurs. Ce constat tend à confirmer notre hypothèse selon laquelle les pays du Sud développent des capacités d'évaluation avant tout pour répondre à une exigence des donateurs et non parce qu'ils ressentent un réel besoin d'évaluation des politiques publiques.

Au-delà de l'existence d'une demande d'évaluation, la littérature relève plusieurs préconditions à la diffusion d'une culture de l'évaluation : la présence d'un État de droit, un bon degré de transparence dans les administrations publiques, un bon niveau d'éducation supérieure avec des universités respectées, des systèmes statistiques solides, un système de production de connaissances abrité du politique, une confiance dans les bénéfices de l'évaluation, une société civile organisée, etc. Ainsi, le renforcement des capacités en évaluation n'est qu'une composante d'un ensemble de réformes de la gouvernance des États[752]. Si l'évaluation est institutionnalisée de manière isolée, sans qu'une réforme plus globale soit

[751] PNUD, 1999, *op. cit.*

[752] HORTON, Douglas, *et al, Evaluation Capacity Development, Experience from Research and Development Organizations around the World,* ISNAR, IDRC, CTA, 2003, 188p, pp. 49-63; BOYLE, Richard, LEMAIRE, Donald, *Building Effective Evaluation Capacity, Lessons from Experience,* Transaction Publishers, New Brunswick, 1999, pp. 135-176, 210p; RIST, Ray, BOILY, Marie-Hélène, MARTIN, Frederick. R, *Development Evaluation in Times of Turbulence, Dealing with Crisis that Endanger our future,* The World Bank, 243p, pp. 143-228.

engagée, il apparaît impossible qu'un système d'évaluation des politiques publiques opérationnel voit le jour[753]. La question qui nous semble alors délicate est précisément de savoir quel est le niveau minimal de « bonne gouvernance » à atteindre pour permettre la mise en place d'un système d'évaluation effectif. Les évaluateurs que nous avons interrogés restent très flous sur la détermination de ce seuil et la littérature n'apporte pas plus de réponses. Toutefois, la pratique montre, comme nous le verrons dans la suite de ce chapitre, que l'institutionnalisation d'un système d'évaluation dans les pays du Sud préexiste souvent à la maturation d'un environnement favorable, ce qui entraîne des difficultés d'opérationnalisation de ces systèmes.

Les donateurs prennent conscience de la nécessité d'agir sur la demande en évaluation et sur l'amélioration du contexte politique et social qui peut la faire naître. Dans la littérature théorique sur le renforcement de la demande en évaluation[754], nous retrouvons trois types de mesures pour augmenter la demande d'évaluation dans les administrations : la « carotte », le « bâton » et le « sermon »[755]. La carotte symbolise une incitation positive qui récompense les activités de suivi et d'évaluation, ainsi que l'utilisation de leurs recommandations. Le « bâton » est un euphémisme pour la mise en œuvre de sanctions légères à l'encontre des acteurs ou organisations qui refusent de pratiquer l'évaluation. Enfin, le « sermon » prend la forme de déclarations des hautes autorités qui sensibilisent les acteurs publics à l'importance de l'évaluation[756].

Nous retrouvons une logique de création d'une demande en évaluation « par le haut », commune aux organismes de développement, selon l'idée qu'il faut, dans un premier temps, forcer la demande en évaluation pour que les fonctionnaires des pays du Sud pratiquent l'évaluation et qu'ils constatent, par cette pratique, son utilité et les bénéfices qu'elle peut leur apporter. Cette stratégie de renforcement de la demande en évaluation se focalise sur les hauts-fonctionnaires et ne vise pas la société civile qui peut pourtant exercer une pression forte sur les gouvernements pour qu'ils adoptent une pratique plus systématique et transparente de l'évaluation, comme ce fut le cas au Maroc[757].

Nous estimons surtout que l'efficacité des actions des donateurs sur la demande en évaluation sera remise en cause tant que les donateurs se substitueront aux pays du Sud dans la mise en œuvre des évaluations des projets et programmes d'aide au développement. Les cas où

[753] Entretien avec le directeur du cabinet d'évaluation Euréval, consultant en évaluation auprès du CAD-OCDE, 20 novembre 2013.
[754] BOYLE, Richard, LEMAIRE, Donald, 2006, *op. cit.* pp. 153-170.
[755] *Ibid*. pp. 153-170.
[756] World Bank, 2006, *op. cit.*
[757] Entretien avec un évaluateur marocain, 17 septembre 2014.

la demande de ces évaluations provient de l'initiative conjointe des donateurs et des bénéficiaires sont encore rares. Les donateurs décident unilatéralement des programmes à évaluer en fonction de leurs besoins en termes de redevabilité ou d'apprentissage interne.

Cette observation nous paraît également vraie en ce qui concerne le renforcement de l'offre en évaluation. Tant que les organismes d'aide privilégieront l'emploi d'évaluateurs du Nord (à 80 % pour les donateurs bilatéraux et à 96 % pour les donateurs multilatéraux[758]), ils nuiront au développement d'une offre nationale d'évaluation, et ce malgré leurs activités de renforcement des capacités. Les donateurs adoptent donc un « double jeu » lorsque, d'une part ils agissent en faveur du renforcement des capacités en évaluation et que, d'autre part ils prennent le pas sur les pays bénéficiaires en matière d'évaluation des programmes d'aide ; ces pratiques s'annulant mutuellement. Selon un évaluateur spécialisé sur la question du renforcement des capacités en évaluation, « *c'est le système mondial de l'évaluation, centré au Nord, dirigé par le Nord, conçu pour répondre aux intérêts du Nord, qui tue la possibilité d'un développement des capacités en évaluation dans les pays en développement. En conséquence, le renforcement des capacités a conduit à des résultats en termes d'augmentation de la connaissance en évaluation, mais très peu en termes de mise en pratique de l'évaluation* »[759]. Cela rejoint notre observation sur les formations en évaluation : si les acteurs du Sud sont de plus en plus nombreux à avoir une expertise théorique en évaluation, ils sont peu à la mettre en pratique.

Ce « double jeu » des donateurs s'explique par leur besoin de garder le contrôle sur le contenu des évaluations, afin qu'elles répondent aux exigences de redevabilité qui pèsent sur eux. Or, les besoins en redevabilité des donateurs (prouver les résultats et l'efficience de leurs programmes d'aide à leurs tutelles financières) sont souvent peu compatibles avec les besoins des pays bénéficiaires (améliorer l'efficacité de la mise en œuvre des politiques publiques, augmenter la transparence de l'administration, apprendre des leçons d'expérimentations économiques et sociales, etc.). Lorsque nous avons interrogé des évaluateurs sur la possibilité de couvrir les besoins des donateurs et des bénéficiaires dans une même évaluation, ils soulignent la difficulté, voire l'impossibilité, d'une telle démarche[760].

Nous identifions deux obstacles principaux : en premier lieu, les évaluations couvrent un nombre d'enjeux important (les cinq critères du CAD) sur des champs de plus en plus larges

[758] CAD-OCDE, 2010, *op. cit.*
[759] Entretien avec le directeur d'un cabinet d'évaluation, consultant en évaluation auprès du CAD-OCDE, 20 novembre 2013.
[760] Entretien avec une évaluatrice de la Banque mondiale, 24 octobre 2013 ; entretien avec une évaluatrice de l'AFD, 25 juin 2014.

(les principes de la Déclaration de Paris/ du Partenariat de Busan et les OMD sur des programmes souvent de grande ampleur, couvrant un secteur d'activité ou une zone géographique élargie). La volonté des donateurs d'avoir des rapports concis et qui apportent des preuves quantitatives de résultats permet difficilement d'avoir une « double perspective », celle des donateurs et celle des bénéficiaires. En second lieu, lorsque les gouvernements bénéficiaires sont intéressés par l'évaluation, il s'agit de l'évaluation de l'efficacité de leurs politiques publiques en général et pas seulement de l'efficacité des programmes des donateurs. Alors que les agences d'aide veulent seulement évaluer leur contribution, les autorités locales souhaitent évaluer l'ensemble des politiques publiques du secteur d'activité concerné. Ce décalage entre les attentes des bailleurs et celles des gouvernements bénéficiaires a été souligné de nombreuses fois par nos interviewés à l'AFD[761].

Les agences d'aide doivent fournir des preuves de redevabilité sur des enjeux de plus en plus divers, ce qui a entraîné un élargissement du nombre et du champ des évaluations. Si les exigences de redevabilité des donateurs étaient plus faibles, les donateurs pourraient mener moins d'évaluations et recourir à des méthodologies qualitatives et plus explicatives permettant un meilleur apprentissage mutuel avec les pays bénéficiaires. Les évaluations seraient alors plus utiles à ces derniers, en leur apportant les clés pour améliorer la gestion de leurs politiques publiques. Toutefois, cela apparaît aujourd'hui difficile alors que le concept de redevabilité est au cœur de l'agenda de l'aide au développement. En témoigne l'attachement des donateurs au respect strict des standards de l'évaluation du CAD-OCDE qui « *constituent le gage de qualité et de redevabilité des évaluations* »[762]. Nos recherches nous ont conduit à penser que si les donateurs sont si attachés au fait que les évaluations traitent rigoureusement chaque critère du CAD et qu'elles mobilisent des méthodologies complexes, sans qu'une réflexion soit menée sur leur pertinence par rapport à l'objet évalué, c'est en raison de leur volonté de se protéger de toute critique de manque de rigueur ou de transparence. Selon un évaluateur des Nations unies, « *les donateurs privilégient l'excellence méthodologique des évaluations par rapport à leur utilité pour les pays du Sud, car leur priorité est d'être redevable et que cette redevabilité repose sur le suivi de ces critères. L'apprentissage par le Sud est un plus mais pas un aspect fondamental de l'évaluation* »[763].

[761] Voir liste des entretiens en annexe 8.
[762] Entretien avec le directeur d'un cabinet d'évaluation, consultant en évaluation auprès du CAD-OCDE, 20 novembre 2013.
[763] Entretien avec un évaluateur des Nations unies, 6 octobre 2014.

Nous observons donc ici, à travers l'opposition entre l'amélioration de la redevabilité des bailleurs de fonds et le renforcement des systèmes nationaux d'évaluation, un exemple des contradictions entre les différentes « bonnes pratiques » de l'agenda de l'aide.

Le renforcement des capacités en évaluation est également affaibli par le fait que ce sont les évaluateurs du Nord, travaillant au sein de, ou pour, des organismes de développement qui réalisent les formations et l'assistance technique en évaluation auprès des pays du Sud. En effet, sur le long terme, un renforcement des capacités en évaluation efficace reviendrait à créer une concurrence nouvelle et à moindre coût aux évaluateurs du Nord. Ces derniers n'ont donc pas intérêt à voir émerger une offre d'évaluation de grande ampleur et structurée dans les pays du Sud. D'après un consultant en évaluation du Nord, « *les évaluateurs du Nord n'ont pas intérêt à promouvoir le développement des capacités en évaluation au Sud. C'est là un des facteurs de l'inefficacité du renforcement des capacités. L'argument des évaluateurs du Nord pour justifier leur manque de coopération avec les experts du Sud est qu'ils ne peuvent pas leur faire confiance en raison de leur appartenance au pays du projet évalué, ce qui nuirait à leur indépendance. Cet argument est peu valable car ces mêmes évaluateurs s'estiment indépendants lorsqu'ils évaluent les politiques publiques de leur État* »[764]. Les évaluateurs du Nord sont, en effet, dans une situation ambiguë. Pour légitimer leur expertise, ils doivent s'associer les compétences d'évaluateurs du Sud selon une logique partenariale de la coopération pour le développement. Mais en même temps, former des évaluateurs compétents au Sud, revient pour eux à se créer de la concurrence. Toutefois, nous estimons que les évaluateurs du Nord sont encore très loin aujourd'hui d'être menacés par la montée d'une expertise au Sud. De plus, les donateurs auront encore longtemps besoin d'une expertise nationale, ou du Nord, pour répondre à des exigences de redevabilité.

Au-delà des formations, le renforcement des capacités en évaluation prend également la forme d'une assistance technique, comme le font, par exemple, la Banque mondiale ou l'agence d'aide DANIDA au Danemark. Dans ce cas, les agences d'aide, appuyées ou non par des consultants externes en évaluation, forment et accompagnent, sur une période plus ou moins longue, le processus de renforcement des capacités en évaluation des pays bénéficiaires en les aidant à diagnostiquer les besoins d'évaluation, à mettre en place des structures, à élaborer des guides méthodologiques, etc. Avec cette approche, l'apprentissage est plus actif et les individus formés appliquent leurs connaissances. Toutefois, nous sommes toujours dans le cadre d'un

[764] Entretien avec le directeur d'un cabinet d'évaluation, consultant en évaluation auprès du CAD-OCDE, 20 novembre 2013.

renforcement des capacités ponctuel qui ne permet pas une réelle appropriation des compétences en évaluation.

Beaucoup d'organismes d'aide se sont inspirés de l'approche de la Banque mondiale et débutent leurs activités de transfert des bonnes pratiques d'évaluation par la recherche d'un « champion » de l'évaluation dans l'administration locale. Toutefois, selon les opérateurs de ce transfert (souvent les évaluateurs des organismes d'aide), il est difficile de trouver de tels « champions » dans des environnements où les compétences et la demande en évaluation sont très faibles. C'est le cas des évaluateurs de l'AFD qui considèrent « la détermination d'un partenaire local soutenant l'évaluation » comme une difficulté majeure du renforcement des capacités[765]. Les autres obstacles au renforcement des capacités en évaluation cités sont le manque de demande et de soutien politique des gouvernements, le manque d'expertise locale, l'insuffisance des ressources financières et des structures allouées à l'évaluation, l'absence d'une bonne appréhension du rôle et du mandat à donner à l'évaluation au sein des administrations, la difficulté d'articuler l'apprentissage théorique et la mise en pratique, ainsi que la difficulté à mobiliser les partenaires locaux sur le long terme[766]. Les problèmes rencontrés par les donateurs sont donc aussi bien politiques que techniques, les premiers étant les plus difficiles à résoudre. Nous comprenons ainsi que l'évaluation possède un caractère politique qui complexifie son importation. Il ne s'agit pas d'un simple instrument technique, dont l'utilisation serait reproductible de manière identique dans toutes les structures organisationnelles et tous les contextes politiques. Les notions sous-jacentes de transparence, d'efficacité des politiques publiques, de redevabilité et d'apprentissage des erreurs passées sont porteuses d'inspirations politiques pour la société civile. L'évaluation est également porteuse des approches de la bonne gouvernance prônées par les donateurs, ce qui peut rendre méfiants les pays bénéficiaires se sentant mis sous la tutelle des donateurs ou craignant une forme d'ingérence de la part de ces derniers. Les « opérateurs de transfert » (les évaluateurs des agences d'aide) de capacités en évaluation témoignent aussi de leur incapacité à effectuer leur tâche, dans la mesure où ils ont une expérience assez limitée en matière d'évaluation. Leur capacité à transmettre leurs compétences et leur expérience est donc limitée, ce qui peut expliquer qu'ils s'en tiennent à la mise en place de formations sur les concepts de base de l'évaluation.

[765] CAD-OCDE, 2006, *op. cit.*
[766] *Ibid.*

Si l'approche de la Banque mondiale en termes de renforcement des capacités en évaluation est la plus reproduite par les organismes d'aide, il n'existe pas de bonnes pratiques institutionnalisées sur cet aspect de l'évaluation. C'est une exception dans un champ où les modèles, standards et guides de procédures se multiplient, aussi bien par secteur (évaluation de la prise en compte du genre, des effets environnementaux, etc.), que par type d'évaluation (évaluations participatives, évaluations conjointes, etc.). Il y a eu des débats au sein des organisations multilatérales pour déterminer s'il était nécessaire ou non d'établir des bonnes pratiques formelles en termes de renforcement des capacités en évaluation. Le principal avantage relevé était que l'existence de bonnes pratiques permettrait de transmettre aux gouvernements bénéficiaires des orientations plus claires sur les standards qu'ils doivent atteindre pour pouvoir considérer qu'ils disposent d'un système d'évaluation satisfaisant. Certains donateurs bilatéraux (le Royaume-Uni, les pays nordiques) estiment que le manque de standardisation et d'harmonisation des pratiques et des objectifs empêche un renforcement des capacités efficace, car les hauts fonctionnaires des pays bénéficiaires se retrouvent exposés à des référentiels normatifs multiples, ce qui empêcherait l'émergence d'une culture harmonisée de l'évaluation au niveau national[767]. De plus, d'après ces donateurs, l'existence de bonnes pratiques permettrait d'évaluer et de comparer plus aisément les systèmes nationaux d'évaluation au Sud.

Cependant, un large consensus existe depuis les années 1990 parmi les évaluateurs qui s'accordent à dire qu'il n'existe pas un modèle universel de bon système d'évaluation[768]. Ils mettent l'accent sur le fait qu'il y a une vaste gamme de modèles envisageables. Cela peut s'expliquer par le fait que l'institutionnalisation de l'évaluation a pris des formes très différentes dans les États du Nord[769]. Si nous comparons, par exemple, le rôle qu'a le gouvernement dans les systèmes d'évaluation, nous constatons de fortes divergences entre les pays. Il est très fort aux États-Unis, au Canada en Suède et en France, alors qu'il est faible au Royaume-Uni, en Suisse et en Allemagne. Concernant le poids de la Cour des comptes ou de son équivalent, il est fort aux États-Unis, au Canada, en Australie et en Suède, tandis qu'il est plus faible en France, au Royaume-Uni et en Allemagne. Enfin, concernant l'influence du Parlement, il est relativement faible partout, excepté aux États-Unis[770].

[767] CAD-OCDE, 2006, *op. cit.*
[768] World Bank, 2011, *op. cit.*
[769] FURUBO, Jean-Eric, RIST, Ray SANDHAL, Rolf, 2002, *op. cit.*
[770] *Ibid.*

Certains chercheurs en politiques publiques et évaluateurs se sont toutefois attachés à théoriser les critères auxquels devrait répondre un système d'évaluation mature[771] :

- l'évaluation prend place dans de nombreux domaines politiques ;
- il existe des évaluateurs compétents dans des disciplines variées ;
- un discours national est développé en matière d'évaluation ;
- des organisations professionnelles d'évaluation ont été créées ;
- l'évaluation est institutionnalisée au sein du gouvernement ;
- l'évaluation est institutionnalisée au sein du parlement ;
- il y a un pluralisme des organisations ou évaluateurs exerçant dans chaque domaine politique ;
- la Cour des comptes (ou son équivalent) pratique l'évaluation ;
- une part significative des évaluations porte sur les résultats des politiques publiques et non sur les processus et les réalisations[772].

Nous tenons, tout d'abord, à souligner qu'il est frappant de constater que l'utilisation et la publicisation des évaluations ne sont pas mentionnées dans ces critères. Cela montre que la culture de l'évaluation a tendance à n'être pensée que d'un point de vue institutionnel et non du point de vue des effets des évaluations. Or, nous estimons que pour parler d'une culture de l'évaluation, il ne suffit pas seulement que des entités administratives réalisent des évaluations, il faut que ces dernières soient utilisées pour améliorer les politiques publiques et/ou pour améliorer le débat public.

Bien qu'il se limite à des aspects institutionnels, ce modèle n'est atteint aujourd'hui que par les États-Unis[773]. Dans ce contexte de grande diversité interétatique, parvenir à un accord international sur un idéal type de système d'évaluation à promouvoir dans l'ensemble des États du Sud apparaît donc difficile, ce qui a conduit les bailleurs de fonds à avoir une approche moins normative que dans d'autres exercices de renforcement des capacités institutionnelles. De plus, l'idéal type décrit ci-dessus n'étant réalisé pleinement que par les États-Unis, les donateurs ne peuvent pas légitimement demander son atteinte par les gouvernements du Sud. Si les organisations internationales reconnaissent qu'il n'y a pas un modèle unique de système d'évaluation efficace, elles donnent, toutefois, des recommandations très similaires aux bénéficiaires. Des donateurs se regroupent régulièrement pour partager leurs expériences en

[771] *Ibid*.
[772] Pour une distinction entre résultats et réalisations voir lexique en annexe 7.
[773] FURUBO, RIST, SANDHAL, 2002, *op. cit*,

matière de renforcement des capacités et pour harmoniser leurs approches[774]. Ce constat nous permet de prendre de la distance par rapport aux propos des donateurs selon qui l'absence de bonnes pratiques formalisées en matière de renforcement des capacités en évaluation ferait que les bénéficiaires seraient confrontés à des référentiels très diversifiés promus par les donateurs.

Selon le CAD-OCDE, pour qu'un système d'évaluation soit opérationnel, il faut *a minima* qu'il existe « une demande en évaluation et que les évaluations soient utilisées durablement et contribuent au processus de prise de décision au niveau pour lequel le système d'évaluation est conçu »[775]. Pour la Banque mondiale, les exigences minimales sont « la création d'un système fournissant des informations de bonne qualité utilisées de manière intensive pour favoriser la bonne gouvernance et qui soit entièrement viable. En d'autres termes, un système capable de survivre à un changement de gouvernement et de continuer à être utilisé par les administrations subséquentes »[776]. La durabilité des systèmes est souvent assimilée à l'inscription de l'évaluation dans la législation. Ainsi, nous observons que les bailleurs de fonds mettent plus en avant la notion d'utilisation des évaluations dans leurs modèles que ne le font les évaluateurs du Nord dont nous avons présenté l'approche ci-dessus.

En effet, l'utilisation des évaluations dans la prise de décision est le critère prouvant l'ancrage d'une culture de l'évaluation le plus cité dans les textes des organismes de développement[777]. Toutefois, nous avons vu dans le chapitre précédent à quel point cette utilisation, souvent diffuse et indirecte, est difficile à appréhender. Même dans les États où la culture de l'évaluation est ancrée dans les politiques publiques depuis plusieurs décennies, les liens entre évaluation et prise de décision restent flous. De plus, c'est spécifiquement l'utilisation des évaluations que les organisations internationales ont le plus de mal à développer dans les pays du Sud – comme c'est d'ailleurs le cas dans les pays du Nord. Dans la pratique, c'est donc plutôt à partir de la mise en place de structures dédiées au suivi-évaluation dans l'administration, de l'existence de sociétés professionnelles d'évaluation et du nombre d'évaluations réalisées par secteur que les capacités en évaluation des pays du Sud sont mesurées.

---

[774] Evaluation Capacity Development Group, *An International Workshop Agreement on Evaluation Capacity Development*, octobre 2011.
[775] SCHAUMBURG-MULLER, Henrik, *Evaluation Capacity Building, Donor Support and Experiences,* Report for the DAC Expert Group on Aid Evaluation, OCDE, 1997.
[776] World Bank, 2006, *op. cit.*
[777] *Ibid ;* SCHAUMBURG-MULLER, 1997, *op. cit ;* UNEG, *National Evaluation Capacity Development: Practical tips on how to strengthen National Evaluation Systems*, 2012.

Nous observons donc que les donateurs prennent de plus en plus conscience du besoin de construire un environnement favorable à l'évaluation, de stimuler l'émergence d'une demande en évaluation et de travailler sur l'utilisation des évaluations autant que sur leur mise en œuvre[778]. Pour ce faire, ils tentent de se tourner aujourd'hui vers un autre vecteur de renforcement des capacités en évaluation : les évaluations partenariales.

B. Les évaluations partenariales, un vecteur d'apprentissage mal exploité

Nous avons observé que la grande majorité des activités en matière de renforcement des capacités en évaluation relève de la formation et de l'assistance technique. Cependant, face aux difficultés évoquées ci-dessus, certains groupes d'évaluateurs, notamment les représentants des pays nordiques au sein du CAD-OCDE, ont commencé à développer des approches alternatives du renforcement des capacités, qui passent par la réalisation d'évaluations partenariales avec les pays du Sud.

Les bénéfices des évaluations partenariales en matière d'apprentissage et de diffusion d'une culture de l'évaluation dans les États du Sud restent aujourd'hui de l'ordre de la théorie. Il n'existe, en effet, pas d'études démontrant à partir d'expériences passées les effets bénéfiques des évaluations partenariales sur le développement de systèmes d'évaluation au Sud. Cette idée que les évaluations partenariales permettraient une diffusion plus effective d'une culture de l'évaluation, que nous partageons, prend ces racines dans les expériences de construction d'une culture de l'évaluation au Nord. Nous avons observé dans le chapitre précédent que le fait que les organismes d'aide au développement fassent participer leurs personnels aux évaluations favorise l'émergence d'une culture organisationnelle de l'évaluation. Nous en avions conclu que l'apprentissage en matière d'évaluation et la diffusion d'une culture de l'évaluation passent par une participation active des acteurs aux processus évaluatifs plus que par des réformes institutionnelles, la publicisation des rapports d'évaluation ou l'apprentissage théorique des bienfaits de l'évaluation.

Nous estimons donc, suivant les théories de Michael Patton[779] et de Carol Weiss[780], que c'est par une participation régulière des membres de l'administration et de la société civile des

[778] UNEDAP, *Regional Thematic Study of national monitoring and evaluation systems in Asia and the Pacific*, Thaïlande, 2012.
[779] PATTON, Michael. Q, 1978, *op. cit.*
[780] WEISS, Carol, 1998, *art. cit.*

pays du Sud aux processus d'évaluation qu'un renforcement des capacités en évaluation a le plus de chance de se produire dans ces pays. En pratiquant l'évaluation (même si cette pratique est financée et encadrée par les donateurs), les acteurs publics du Sud devraient mieux s'approprier les processus évaluatifs, que s'ils les abordaient par la théorie. Convaincus par les bénéfices de la démarche évaluative, ils seraient plus à même ensuite de sensibiliser leurs administrations à l'évaluation et de les inciter à développer un système d'évaluation. Faire participer les bénéficiaires des programmes d'aide aux évaluations permettrait ainsi d'agir à la fois sur l'offre et la demande d'évaluation.

En effet, si nous reprenons la typologie d'Harold Wolman[781], dans le cadre de formations ou de séminaires, les représentants des pays du Sud ne sont que les récepteurs de la compétence transmise. Ils ne deviendront opérateurs ou facilitateurs que s'ils appliquent ces compétences dans leur pays d'origine.

Dans le cadre d'évaluations partenariales, le positionnement des représentants des pays du Sud change. Ils deviennent « opérateurs », au même titre que les fonctionnaires ou experts internationaux, car ils prennent une part active au transfert. La relation devient ainsi plus égalitaire entre exportateur et importateur de la compétence. De plus, le *learning by doing* permis par les évaluations partenariales semble pour beaucoup plus approprié que la formation théorique pour transférer des compétences en évaluation. Un évaluateur explique que « *l'évaluation est un métier d'artisan. Un métier qui s'apprend par la pratique. L'évaluation partenariale est le meilleur moyen de formation. La coproduction d'évaluations permet un apprentissage mutuel entre l'évaluateur du Nord et du Sud* »[782].

Nous allons désormais analyser les effets concrets des évaluations partenariales en matière de renforcement des capacités en évaluation, afin de comprendre dans quelle mesure et à quelles conditions elles peuvent constituer des vecteurs de diffusion d'une culture de l'évaluation dans les pays du Sud.

- **Des partenariats déséquilibrés**

En 2013, le réseau d'évaluateurs du CAD-OCDE et la Société française d'évaluation (SFE) ont développé un axe de travail sur le rôle des évaluations partenariales dans la création

[781] WOLMAN, 2002, *ort. cit.*
[782] Entretien avec le directeur d'un cabinet d'évaluation, 9 décembre 2014.

d'une offre et d'une demande d'évaluation au Sud. Claude Leroy Thémèze, évaluatrice à la Banque mondiale, et Jacques Toulemonde, directeur retraité d'un cabinet français d'évaluation et consultant auprès du CAD-OCDE, sont à la tête de cette initiative.

En France, une initiative similaire de promotion du renforcement des capacités en évaluation au Sud par l'évaluation partenariale avait été mise en place par le ministère des affaires étrangères. Selon un évaluateur ayant participé à l'évaluation du bilan de cette expérimentation, elle fut un échec[783]. Lorsque nous avons été associée au projet d'Evalnet et de la SFE, cette expérimentation n'a jamais été mentionnée et les leçons de cet échec ne semblent pas avoir été partagées entre le MAE, la DG Trésor et la SFE, ce qui témoigne de la faible communication, voire de la concurrence, entre ces entités d'évaluation. C'est donc sur une base nouvelle que le projet a été lancé.

Claude Leroy-Thémèze estime que « *le learning by doing passe par l'implication dans les évaluations menées par les donateurs de partenaires locaux qui vont apprendre de leur participation. Les institutions du pays partenaires vont ainsi être incitées à modifier leurs routines et leurs pratiques et à utiliser et renforcer les systèmes d'évaluation nationaux qui vont ensuite être le vecteur de changements culturels. Ces systèmes vont ainsi être légitimés et renforcés* »[784]. Cette approche du renforcement des capacités en évaluation renverse la conception habituelle de l'apprentissage, selon laquelle on part d'une accumulation de connaissances théoriques que l'on met ensuite en pratique. Il s'agit plutôt ici d'envisager un modèle d'apprentissage dans lequel on part de la pratique pour tirer ensuite des connaissances de cette expérience. Cette approche joue donc à la fois sur la demande, l'offre et l'utilisation et entend ainsi renforcer l'ensemble de la culture de l'évaluation, afin d'obtenir, selon les termes de Jacques Toulemonde : « *une demande durable d'évaluation à des niveaux et dans des secteurs divers ; une offre de services d'évaluation compétents ; l'utilisation des résultats d'évaluation dans le secteur public ; un cadre institutionnel permettant une confiance publique en l'évaluation* »[785]. Les tenants de l'approche du renforcement des capacités en évaluation par les évaluations partenariales considèrent qu'il est indispensable que les partenaires locaux soient impliqués le plus largement possible à toutes les étapes des évaluations des programmes d'aide pour qu'il y ait un renforcement des capacités efficaces, chose qui n'a encore jamais réellement été faite.

[783] Le ministère des affaires étrangères n'a pas souhaité nous communiquer plus de détails sur les raisons et la nature de l'échec de cette expérimentation.
[784] Propos recueillis lors d'un séminaire de travail du réseau d'évaluateurs impliqués dans cette initiative.
[785] *Ibid.*

Selon nous, toute la difficulté de la mise en place de ce type d'apprentissage est de rompre avec le système actuel, dans lequel l'évaluation de l'aide est faite par et pour les donateurs. Ce système implique des conséquences négatives nombreuses sur le renforcement des capacités en évaluation au Sud [786] :

- il affaiblit la légitimité des réseaux nationaux et des mécanismes d'évaluation en les détruisant indirectement ;
- il empêche l'établissement de mécanismes d'évaluation dans le pays partenaire lorsque les acteurs et organisations se reposent trop sur les règles, guides et standards des donateurs ;
- il fait diverger les activités d'évaluation par rapport aux besoins du pays partenaire en se concentrant trop sur la redevabilité des pays partenaires ;
- il n'implique pas les partenaires dans les tâches centrales de l'évaluation : la rédaction des termes de référence, l'élaboration de la méthodologie, la formulation des recommandations ;
- il surcharge les systèmes nationaux par la conduite simultanée de nombreuses évaluations ;
- il distord le marché de l'évaluation en établissant des salaires et tarifs supérieurs à la moyenne nationale ;
- il échoue à renforcer les organisations nationales en se reposant sur des évaluateurs en free-lance ou sur des évaluateurs des pays donateurs.

Les membres du réseau Evalnet du CAD-OCDE sont conscients de l'impossibilité de mettre en œuvre, dans le système actuel de l'aide au développement, des évaluations parfaitement partenariales, dans lesquelles les pays bénéficiaires auraient le même poids d'analyse et de décision que les donateurs[787]. C'est pour cela que leur première démarche a été de lancer des études consistant à « identifier à partir d'expériences concrètes récentes, les moyens de promouvoir la pratique des évaluations partenariales avec les pays bénéficiaires, qui favorisent le recours aux évaluateurs du Sud et le renforcement des capacités d'évaluations dans les pays bénéficiaires », afin de « faire comprendre aux donateurs et aux bénéficiaires le potentiel des évaluations partenariales en matière de renforcement des capacités, tirer des

[786] DAC-OCDE, *Study on Collaborative Partner-Donor Evaluation Work*, Concept Note, document de travail, Evalnet, 26 juin 2013.
[787] Nous avons pu faire ce constat lors de notre participation à des réunions du groupe Evalnet à Paris en 2013.

nouvelles leçons à partir des stratégies de renforcement des capacités et mener à une meilleure collaboration entre les donateurs et les bénéficiaires »[788].

La première difficulté pour analyser les effets d'apprentissage des évaluations partenariales consiste à trouver des cas d'étude permettant de tirer des conclusions sur les liens entre évaluations partenariales et renforcement des capacités. En effet, les évaluations réellement partenariales, avec une implication significative des pays partenaires sont très rares. D'après nos recherches, l'agence danoise d'aide au développement (DANIDA), semble être, de loin, l'agence d'aide la plus largement investie dans la mise en place d'évaluations partenariales. Cette organisation est considérée par les autres donateurs comme un modèle en la matière[789]. Nous allons donc partir de l'expérience de cette agence pour déterminer le degré d'influence que peuvent avoir, dans le contexte actuel, les évaluations partenariales sur le renforcement des capacités des pays du Sud.

La DANIDA a mené dix évaluations partenariales depuis le début des années 2000, soit de niveau sectoriel, soit de niveau national[790]. Il faut tout d'abord noter que ces évaluations n'avaient pas pour objectif spécifique le renforcement des capacités en évaluation des pays au sein desquels les programmes évalués avaient été mis en œuvre. Le renforcement des capacités recherché se limitait aux consultants locaux impliqués dans l'évaluation. S'il y a eu renforcement des capacités à une plus large échelle, il s'agit donc d'un effet indirect de ces évaluations partenariales[791].

Dans la totalité des évaluations, la demande provenait de la DANIDA. Les partenaires béninois, ghanéens et mozambicains ayant participé aux évaluations conjointes ont estimé qu'ils auraient dû être impliqués plus en amont du processus évaluatif. L'un d'entre eux témoigne par exemple : « l'évaluation nous est tombée dessus sans que l'on ne nous prévienne »[792]. Ils évoquent également le fait que la préparation des termes de référence était déjà trop avancée lorsqu'ils ont été impliqués dans le processus évaluatif pour qu'ils puissent

[788] Evalnet, juin 2013, *op. cit.*
[789] Constat issu de discussions informelles avec les évaluateurs de l'AFD lors de notre observation participante dans cette organisation en 2012 et 2013.
[790] Joint Evaluation of The Danish Country Programme in Benin 1990-2006, 2008; Joint evaluation of the Ghana-Denmark Development Cooperation 1990-2006, 2008; Evaluation of Development Cooperation between Mozambique and Denmark 1992-2006, 2008; Evaluation of Danish Assistance to Uganda 1987-2005, 2006; Capacity Development Outcome Evaluation of Danish Supported Organizations in Bhutan, 2006; Joint Evaluation of the Road sub-sector Programme 1996-2000 in Ghana, 2000; Joint Government-Donor Evaluation of Basic and Primary Education in Nepal, 2004; Joint External Evaluation, the Health Sector in Tanzania, 1999-2006, 2007; Joint Evaluation of Uganda's Plan for the Modernization of Agriculture, 2005.
[791] Nous pouvons donc penser que dans le cas d'évaluations partenariales qui seraient mises en place dans l'optique première de renforcer les capacités de l'administration d'un pays en évaluation les résultats trouvés seraient meilleurs que ceux observés dans notre étude.
[792] Ministry of Foreign Affairs of Denmark, *Experience with conducting evaluation jointly with partner countries*, Danida's Evaluation Department, juin 2009.

apporter de changements significatifs sur l'approche adoptée ou les enjeux traités. Ces initiatives unilatérales d'évaluations partenariales ont parfois été mal accueillies par les États partenaires, et ce pour deux principales raisons : le manque de valeur ajoutée perçu de l'évaluation lorsqu'elle faisait suite à de nombreuses études effectuées par d'autres donateurs dans le secteur évalué, ou bien une mauvaise temporalité de l'évaluation dans l'agenda politique national. Les évaluateurs du Nord justifient cette faible association des partenaires locaux dans les premières phases des évaluations par le fait que le temps pour répondre aux appels d'offres est très court et qu'ils n'ont pas le temps de rechercher et d'associer des experts des pays du Sud à cette phase, d'autant plus que l'expertise disponible leur est peu connue. Ces constats sur l'implication des États partenaires dans la commande et la préparation des évaluations tendent à montrer que les pratiques actuelles de la DANIDA, qui rejoignent celles des autres donateurs, ne permettent pas un renforcement des capacités sur ces aspects du processus évaluatif[793].

Ensuite, un « groupe de référence »[794] est établi, constitué de représentants du pays donateur et du pays bénéficiaire. Deux stratégies existent quant au choix des organisations à impliquer dans la mise en œuvre d'une évaluation partenariale. Dans la plupart des cas, la DANIDA a choisi comme organisation partenaire un ministère stratégique « englobant », comme le ministère des finances, le ministère du développement, ou encore de la planification. Cette stratégie rejoint celle du « champion » de la Banque mondiale, qui consiste à trouver une personne haut placée dans l'administration ayant un réel intérêt pour l'évaluation et à la former pour qu'elle impulse par la suite une culture de l'évaluation au sein de l'administration. Elle permet également de renforcer l'indépendance et la crédibilité perçues de l'évaluation, ce qui rentre dans l'objectif du donateur d'être le plus redevable, et donc indépendant, possible. L'autre stratégie consiste à prendre comme organisation partenaire le ministère sectoriel correspondant au programme évalué. Les évaluateurs de la DANIDA considèrent que ce choix favorise l'appropriation et l'utilisation des évaluations, mais qu'il n'est pas facile à assumer, car il fragilise leur indépendance[795]. Nous retrouvons, encore une fois, le dilemme de l'évaluateur entre appropriation locale et redevabilité. Dans tous les cas, le choix du partenaire de l'évaluation n'est jamais précédé d'un diagnostic des capacités en évaluation des administrations du pays. La raison donnée par les évaluateurs de la DANIDA est que les capacités en évaluation des États sont tellement faibles que ce ne peut être cette caractéristique qui détermine le choix de l'organisation partenaire. Ce choix est davantage déterminé par la

[793] DAC-OECD, 2010, *op. cit.*
[794] Voir lexique en annexe 7.
[795] Ministry of Foreign Affairs of Denmark, 2009, *op. cit.*

pertinence du positionnement institutionnel de la structure partenaire par rapport à la finalité recherchée de l'évaluation[796].

Une fois le cadre de l'évaluation déterminé, le donateur et le pays partenaire sélectionnent une équipe d'évaluateurs. Dans chaque évaluation partenariale étudiée, l'équipe d'évaluateurs est composée de consultants danois et de consultants locaux. Alors qu'au Nord, des cabinets de consultants sont recrutés, il s'agit au Sud, soit de consultants en free-lance, les regroupements d'évaluateurs dans des structures formelles étant rares, soit de fonctionnaires haut placés (membre du cabinet du premier ministre, directrice de la commission nationale pour le développement, directeur adjoint du ministère de la planification). L'équilibre entre le nombre de consultants du Nord et du Sud varie selon les évaluations.

Au Ghana et au Bénin, nous retrouvons les mêmes évaluateurs du Sud dans les évaluations de la DANIDA et dans celles d'autres donateurs, notamment l'évaluation partenariale de la Déclaration de Paris, qui sera étudiée ci-dessous, ou encore des évaluations de l'AFD. Il semble donc que certains acteurs du Sud se spécialisent dans ce type d'exercice et deviennent la référence locale en évaluation des donateurs[797]. Un interviewé chinois témoigne ainsi « *qu'après avoir participé à une évaluation partenariale avec l'agence norvégienne NORAD, il a été invité par des donateurs français et hollandais à participer à des exercices d'évaluations partenariales* »[798]. Le rôle de ces acteurs est difficile à déterminer car, selon les évaluations, ils participent soient en tant que membre du comité de pilotage, soit en tant qu'évaluateur.

Dans la moitié des évaluations étudiées, nous trouvons que la collaboration avec le pays partenaire est déséquilibrée. Les évaluateurs du Nord ont la responsabilité des phases clés de l'évaluation (collecte de données, analyse des données, formulation des recommandations) et les évaluateurs du Sud ont un rôle de facilitateur se limitant à des aspects logistiques (organisation de missions de terrain, d'ateliers de travail ou de réunions)[799]. L'analyse des processus évaluatifs fait donc ressortir une faible implication des évaluateurs du Sud dans les aspects de planification, de production et d'analyse des données et de formulation des

[796] *Ibid.*
[797] Notamment, Justine A. Odjoubé, Directrice de l'observatoire du Changement Social, Ministère d'État Chargé de la Prospective, du *Développement et de l'Évaluation de l'Action Publique* qui participe activement aux évaluations partenariales des donateurs, aux séminaires internationaux qu'ils organisent sur l'évaluation et aux réseaux internationaux d'évaluateurs.
[798] Entretien avec l'Executive Director du National Center for Science and Technology Evaluation (NCSTE), Chine, 8 juin 2013.
[799] Evaluation of Danish Assistance to Uganda 1987-2005, 2006; Capacity Development Outcome Evaluation of Danish Supported Organizations in Bhutan, 2006; Joint Government-Donor Evaluation of Basic and Primary Education in Nepal, 2004; Joint External Evaluation, the Health Sector in Tanzania, 1999-2006, 2007; Joint Evaluation of Uganda's Plan for the Modernization of Agriculture, 2005.

recommandations, qui sont justement les éléments du processus évaluatif les plus porteurs d'un potentiel effet d'apprentissage. Lorsque nous avons interrogé les acteurs du Sud et les membres du CAD-OCDE sur les raisons de cette situation, plusieurs arguments communs ressortent de leurs discours : « *la barrière de la langue* », « *les difficultés de communication entre les évaluateurs et les acteurs locaux, principalement dues au vocabulaire très technique utilisé par les évaluateurs du Nord* », « *la trop grande complexité des méthodologies utilisées* » et « *les références théoriques trop pointues mobilisées par les évaluateurs du Nord* »[800]. Toutefois, les évaluateurs du Nord ont une autre lecture du problème : « *si les consultants du Sud sont placés dans des positions de facilitateurs, d'appui logistique ou de fournisseurs de carnets d'adresses, c'est parce qu'ils ne souhaitent pas s'impliquer davantage dans les processus évaluatifs. En effet, quelle que soit leur contribution, ils touchent la même rémunération. Ils ne sont alors pas incités à s'engager dans le travail plus exigeant qu'est la production de données ou la rédaction des rapports. Il est très rare de voir des consultants du Sud vouloir s'investir dans ces tâches* »[801].

Au-delà du trop faible degré de participation des partenaires du Sud, nous avons identifié d'autres facteurs allant à l'encontre du renforcement des capacités en évaluation : la faible pertinence des évaluations pour les partenaires locaux, aussi bien dans leur contenu, que dans leur temporalité, la mise en œuvre trop ponctuelle d'évaluations partenariales et le fait que le renforcement des capacités en évaluation ne soit la priorité ni du donateur, ni du pays partenaire[802].

En conséquence, les effets d'apprentissage dépendent de l'intensité de la volonté des évaluateurs du Sud à se mobiliser pour tirer eux-mêmes des leçons de l'expérience des évaluations partenariales qu'ils ont vécues, l'organisation de ces évaluations ne favorisant pas en elle-même l'apprentissage. Généralement, les évaluateurs du Sud déclarent avoir gagné en expérience lors des évaluations partenariales, mais leur savoir ne se diffuse qu'à la condition qu'ils le partagent avec d'autres par la suite. Les officiels locaux impliqués dans les évaluations font des retours plus mitigés sur les compétences qu'ils ont pu acquérir à partir de leur expérience[803]. Par exemple, dans le cas ghanéen, ils déclarent « *avoir gagné une expérience utile* », alors que dans le cas béninois, ils témoignent « *n'avoir pas amélioré leurs compétences*

[800] Arguments recueillis à l'issue d'un séminaire réalisée par la DG Trésor et la Société française d'évaluation sur l'implication des pays du Sud dans les évaluations des programmes d'aide au développement. Voir, liste des évènements observés en annexe 9.
[801] Entretien avec le directeur d'un cabinet d'évaluation des politiques publiques basé en France, 9 décembre 2014.
[802] *Ibid*.
[803] DAC-OECD (b), 2013, *op. cit.*

en évaluation, ni leur connaissance sur le sujet évalué ». Dans tous les cas, même lorsqu'il y a un apprentissage individuel, il ne se traduit pas par un renforcement de la culture de l'évaluation au niveau organisationnel. Justine Odjoube, haut fonctionnaire béninoise ayant participé à de nombreuses évaluations partenariales, déclare à propos d'une expérience avec l'AFD : « *aucun transfert de connaissances n'a eu lieu au niveau de l'administration. Le consultant béninois a fait son travail tout seul et l'a transmis pour approbation à l'AFD sans chercher à partager les compétences techniques qu'il avait pu acquérir* »[804].

De cette analyse, il ressort que les évaluations partenariales menées par la DANIDA ont eu peu d'effets en termes de renforcement des capacités en évaluation au-delà du niveau individuel. Pourtant, la DANIDA est une des agences d'aide qui s'engage le plus dans l'implication des parties prenantes du Sud à ses évaluations. Pour parvenir à ce que les évaluations partenariales soient un vecteur de renforcement des capacités en évaluation au Sud, il paraît nécessaire de faire de cette finalité un objectif premier. Il semble également indispensable de parvenir à modifier en profondeur les fondements du système international de l'évaluation qui reste aujourd'hui centré au Nord. Même dans les évaluations considérées comme les plus partenariales, des caractéristiques nuisibles au renforcement des capacités persistent, comme la mainmise des organismes du Nord sur les phases stratégiques de l'évaluation, la non-utilisation des canaux de financement et d'approvisionnement locaux, ou la volonté des évaluateurs du Nord d'utiliser des méthodologies très complexes hors de portée d'acteurs débutant dans l'évaluation. Du côté des acteurs du Sud, il manque un réel intérêt pour l'évaluation, renforcé par le sentiment que les évaluations des programmes d'aide ne servent que les intérêts des donateurs. Il y a un défaut d'incitations pour que les évaluateurs du Sud s'impliquent davantage dans les phases techniques des évaluations.

Nous allons désormais étudier une expérimentation d'évaluation partenariale d'échelle internationale, la première jamais réalisée, dont un des objectifs principaux était justement un renforcement global des capacités en évaluation dans les pays du Sud. Il s'agit de l'évaluation de la mise en œuvre des principes de la Déclaration de Paris réalisée par le CAD-OCDE en 2011. Cette évaluation était censée constituer un modèle d'évaluation partenariale avec une intégration extensive et complète des pays partenaires. Nous allons analyser si elle est parvenue à dépasser les obstacles des évaluations partenariales que nous avons identifiés ci-dessus.

[804] Entretien avec la directrice de l'Observatoire du changement social au Bénin, 2 décembre 2013.

- **L'évaluation de la Déclaration de Paris : un modèle de partenariat ?**

Pour examiner les progrès réalisés par les organismes d'aide et les pays du Sud dans l'application des principes de la Déclaration de Paris, le CAD-OCDE a mis en place un mécanisme de suivi-évaluation de grande ampleur. À travers ce mécanisme, le CAD voulait montrer qu'il était possible de mener des évaluations à la fois partenariales et rigoureuses de l'aide au développement, et ce à l'échelle internationale. Tout d'abord, des enquêtes de suivi ont été réalisées en 2006, 2008 et 2011. Elles visaient à fournir des « clichés instantanés » de l'évolution des pratiques des donateurs et de l'efficacité de l'aide au vu des indicateurs et des objectifs de la Déclaration de Paris[805]. Les enquêtes ont été organisées dans une quarantaine de pays du Sud, ce qui a permis au CAD-OCDE d'étudier l'application des principes de la Déclaration de Paris sur le terrain, ainsi que de recueillir le point de vue des partenaires locaux et de mobiliser leur expertise dans une démarche participative.

Ces enquêtes de suivi ont servi de base à l'évaluation de la Déclaration de Paris, réalisée en deux phases en 2007 et en 2011[806]. L'évaluation a été conçue pour regrouper l'ensemble des « bonnes pratiques » en matière d'évaluation promues par l'agenda sur l'efficacité de l'aide[807], à savoir la participation, la transparence, l'indépendance, la pertinence, l'utilité, la participation, la redevabilité mutuelle, l'appréciation de l'efficience et de l'efficacité et la rigueur. Cela a été la plus grande évaluation partenariale jamais réalisée : 11 donateurs et huit pays partenaires ont participé à la première phase et 32 pays partenaires et 18 donateurs à la seconde. L'évaluation avait pour objectif d'examiner la performance globale des donateurs et des bénéficiaires de l'aide dans la mise en œuvre de la Déclaration de Paris. Elle visait également à analyser la contribution qu'a eue la Déclaration à l'amélioration de l'efficacité de l'aide. Enfin, elle avait l'ambition de renforcer les capacités en évaluation des pays du Sud en les intégrant pleinement au processus évaluatif.

Étant donné que cette évaluation, d'une part s'est posée en modèle de l'évaluation partenariale, et, d'autre part a constitué le point de départ de l'engagement du réseau Evalnet du CAD-OCDE pour le renforcement des capacités en évaluation par le *learning by doing,* il nous est apparu intéressant de revenir sur sa mise en œuvre et sur ses résultats en termes d'apprentissage pour les pays du Sud. Nous allons analyser quelles opportunités et difficultés ont rencontré les participants à l'évaluation et dans quelle mesure elle a pu contribuer à la

[805]DAC-OECD, 2005, *op. cit.*
[806] WOOD et al, *L'évaluation de la Déclaration de Paris, Phase 2 rapport final,* DAC-OECD, mai 2011.
[807] L'agenda sur l'efficacité de l'aide regroupe tous les engagements pris par les donateurs et les pays bénéficiaires de l'aide dans les « hauts sommets » sur l'efficacité de l'aide, de la Déclaration de Rome au Partenariat de Busan.

diffusion d'une culture de l'évaluation au Sud. Afin d'avoir la vision la plus juste possible de l'expérience qu'ont eue les parties prenantes de l'évaluation, nous avons basé notre réflexion sur les témoignages de participants du Nord et du Sud[808], en partie retranscrits et analysés ci-dessous.

Concernant le regard des évaluateurs du CAD (donc les commanditaires de l'évaluation) sur cette expérience, Dorte Kabell[809], consultante en évaluation au Danemark, chargée de la coordination et de la synthèse des « rapports pays » de l'évaluation, porte un regard assez critique sur le partenariat entre le Nord et le Sud qui a eu lieu durant l'évaluation : « lors de la première phase d'évaluation (en 2007), le constat a été que les interactions entre les parties prenantes du Nord et du Sud avaient été limitées. Un effort particulier a été fait sur cet aspect durant la seconde phase avec l'organisation d'ateliers d'étude pour assurer que le processus d'évaluation était bien compris par les différents acteurs de l'évaluation ». Elle identifie plusieurs facteurs ayant rendu difficiles le dialogue et le partenariat entre les acteurs du Nord et du Sud et donc la concrétisation d'un apprentissage mutuel. Tout d'abord, « la méthodologie de l'évaluation était trop complexe pour être appropriée rapidement par des acteurs ayant peu d'expérience dans cette discipline ». En effet, elle explique que « *les concepteurs de l'évaluation avaient fait ce choix d'une méthodologie complexe, car ils craignaient de se faire critiquer sur le manque de rigueur et de qualité de l'évaluation* ». De plus, au vu du nombre important de parties prenantes impliquées, « *le temps de l'évaluation était trop court pour permettre une participation substantielle des acteurs et un dialogue approfondi sur les réalisations de la Déclaration de Paris* ». Le temps était d'autant plus court que les dispositifs de suivi et d'évaluation étaient nombreux (trois enquêtes de suivi et deux évaluations).

En effet, selon nous, l'objectif donné à l'évaluation de mesurer la contribution de la Déclaration de Paris à une éventuelle hausse de l'efficacité de l'aide au développement a nui à l'objectif de renforcement des capacités des partenaires du Sud. Nous percevons ici le dilemme entre le besoin des donateurs d'être le plus redevables possible, qui passe par l'utilisation de méthodologies complexes d'évaluation, et l'engagement éthique de mener les évaluations de manière partenariale avec les pays du Sud. Pourtant, si certaines parties prenantes de l'évaluation soulignent sa complexité, l'évaluation de la Déclaration de Paris a privilégié des

[808] Ces témoignages ont pu être recueillis lors de l'organisation et du déroulement d'une conférence intitulée : *l'évaluation de l'aide au développement, quelles perspectives ?* réalisée conjointement par l'unité d'évaluation du Trésor français et la Société française d'évaluation. Cette conférence faisait interagir des acteurs du développement du Nord et du Sud ayant participé à l'évaluation de la Déclaration de Paris.

[809] Avant de monter son cabinet de conseil évaluation au Danemark, Dorte Kabell a été *Senior Advisor to the Vice President* de la Banque asiatique de Développement, *Senior Addvisor to the President* de la Banque africaine de Développement et *administrator* au CAD-OCDE sur la gestion de l'aide au développement et l'environnement.

méthodes qualitatives d'analyse en raison du manque de données quantitatives disponibles dans les pays du Sud révélé par les enquêtes de suivi, ce qui a été critiqué comme un manque de fiabilité par certains[810]. L'équipe d'évaluation a « compensé » ce manque par un arsenal lourd de collecte de données qualitatives qu'il a été difficile de mobiliser au vu de la courte durée du processus d'évaluation et du nombre de pays évalués. Un autre consultant engagé pour coordonner l'évaluation dévoile à ce sujet que « *le processus de suivi-évaluation a été une « arnaque » sur plusieurs points. Deux enquêtes de suivi auraient dû nourrir les deux phases de l'évaluation*[811]*. Lors de la première phase, l'évaluation a été faite en catastrophe entre la réception de l'enquête de suivi et le sommet d'Accra ; le temps était trop court. Pour éviter cela dans la seconde phase, l'évaluation a été faite de manière plus anticipée que prévu. Elle a été rendue huit mois avant le sommet de Busan, alors que l'enquête de suivi n'était pas terminée. Les évaluateurs ont donc utilisé de nouveau la première enquête. Il n'y avait aucune nouvelle donnée chiffrée. La deuxième évaluation n'a, en réalité, été qu'un affinage qualitatif de la première. De plus, cette seconde évaluation n'a jamais été prise en compte dans la rédaction du rapport de Busan* »[812]. La surenchère d'outils de suivi-évaluation dans un temps très court a donc été contreproductive et a affaibli l'apprentissage permis par l'évaluation. Nous comprenons ainsi que la technocratisation amoindrit l'apprentissage et plus largement la possibilité d'une collaboration équilibrée entre les évaluateurs du Nord et du Sud. C'est d'autant plus regrettable pour les parties prenantes de l'évaluation de la Déclaration de Paris que celle-ci n'est finalement pas parvenue à apprécier la valeur ajoutée de la mise en œuvre des bonnes pratiques de la Déclaration de Paris sur l'efficacité générale de l'aide au développement. Il y a donc eu une double perte, d'une part en matière de mesure de l'efficacité et de redevabilité et, d'autre part en matière d'apprentissage mutuel et d'amélioration du partenariat Nord-Sud[813].

Marjolaine Nicod[814], coordinatrice du CAD-OCDE, fait également part de difficultés de communication avec certains représentants de gouvernements de pays du Sud, peu « *sensibilisés à l'éthique de l'évaluation* » et qui auraient exercé « *une forte pression politique pour que les conclusions des évaluations soient modifiées* ». Ce constat aurait amené les membres du CAD à prendre conscience, qu'au-delà du renforcement des capacités techniques, les formations en évaluation dans les pays du Sud devraient davantage aborder les aspects

[810] BERTHELEMY, Jean-Claude, Sur la déclaration de Paris, Évaluer l'évaluation, document de travail 41, FERDI, mars 2012.
[811] En plus de l'enquête de base en 2006 pour constituer une référence de départ à l'évaluation des progrès réalisés.
[812] Discussion informelle avec un évaluateur de l'équipe de coordination de l'évaluation réalisé à la suite de la conférence du Trésor et de la Société française d'évaluation susmentionnée.
[813] BERTHELEMY, Jean-Claude, *op. cit.*
[814] Marjolaine Nicod a fait toute sa carrière (20 ans) au CAD-OCDE, elle fut entrecoupée par un passage au bureau du DFID au Cambodge durant deux ans.

éthiques et politiques de l'évaluation. Par ailleurs, les coordinateurs de l'évaluation ont constaté, lors de l'exercice de synthèse des « rapports pays », que la qualité des évaluations variait grandement entre les pays, aussi bien au Nord qu'au Sud. L'un des coordinateurs déclare : « *dans les pays d'Afrique francophone, les rapports sont, soit d'une qualité très décevante, soit réellement bien faits* »[815]. L'expérience de l'évaluation de la Déclaration de Paris a donc permis de mieux percevoir les besoins des pays du Sud en matière de renforcement des capacités en évaluation et les diversités des compétences dans ces pays.

Les témoignages recueillis montrent que les difficultés inhérentes aux évaluations participatives sont encore mal maîtrisées par les évaluateurs du Nord, notamment lorsqu'elles s'accumulent aux exigences de la redevabilité. Ils révèlent aussi que l'affichage d'une redevabilité et de pratiques exemplaires n'empêche pas l'opacité des processus d'évaluation. En effet, les aléas évoqués ci-dessus concernant la non prise en compte de certaines enquêtes de suivi dans l'évaluation finale n'ont jamais été divulgués dans les allocutions des représentants du CAD-OCDE ou dans les documents officiels de l'organisation.

Malgré les difficultés éprouvées, les évaluateurs du Nord que nous avons interrogés estiment que l'évaluation de la Déclaration de Paris a eu plusieurs effets positifs sur la sensibilisation des donateurs à l'évaluation partenariale et sur le renforcement des capacités des pays du Sud[816]. Selon eux, le travail participatif qui a eu lieu, même imparfait, a permis une meilleure compréhension par les parties prenantes des contraintes des uns et des autres, ainsi qu'une augmentation de la confiance mutuelle. Elle a également entraîné, au Nord comme au Sud, l'acquisition de nouvelles compétences techniques en évaluation. Un membre de l'équipe d'évaluateurs du Nord déclare par exemple : « *l'évaluation de la Déclaration de Paris fut un très bel exercice pour discuter avec les pays. Des choses très différentes de ce que l'on attendait sont ressorties du dialogue avec les bénéficiaires* »[817], ou encore un membre du réseau d'évaluateurs du CAD-OCDE : « *nous avons des exemples d'évaluations partenariales réussies, comme par exemple celle de la Déclaration de Paris qui s'est déroulée de manière indépendante et participative* »[818]. Dans les allocutions et les rapports publics, c'est l'image d'une évaluation exemplaire en termes de partenariat équitable et d'apprentissage mutuel qui

[815] Entretien avec un consultant en évaluation, 23 mars 2013.
[816] Discussion informelle à l'issue Discussion informelle avec un évaluateur de l'équipe de coordination de l'évaluation réalisé à la suite de la conférence du Trésor et de la Société française d'évaluation susmentionnée. ; entretien avec une Senior Policy Advisor au sein du département de l'efficacité de l'aide du CAD-OCDE, 2 juillet 2014 ; entretien avec un ancien évaluateur du ministère de la coopération français, consultant auprès d'organisations internationales, 21 mars 2013.
[817] Entretien avec un ancien évaluateur du ministère de la coopération français, consultant auprès d'organisations internationales, 21 mars 2013.
[818] Entretien avec un membre du réseau d'évaluateurs du CAD-OCDE, entretien à Paris, 26 novembre 2012.

est unanimement diffusée par les acteurs de l'évaluation. Toutefois, cette communication très positive autour de la Déclaration de Paris cache, comme nous l'avons montré, des difficultés importantes de mise en œuvre qui apparaissent de manière encore plus sensible lorsque nous interrogeons les participants du Sud à l'évaluation.

Pour recueillir les retours d'expérience des parties prenantes du Sud, nous nous sommes appuyée sur les témoignages de membres béninois et sénégalais de l'équipe d'évaluation de la Déclaration de Paris[819] et de hauts fonctionnaires de ces pays qui ont été associés à l'exercice. En premier lieu, nous tenons à souligner qu'un élément particulièrement intéressant de ces témoignages est qu'ils reproduisent le discours technocratique des évaluateurs. À travers les récits des parties prenantes du Sud, nous percevons l'influence qu'a eue cette expérience sur leur perception de l'évaluation.

Les aspects positifs et négatifs de l'évaluation en matière de renforcement des capacités relevés par les intervenants du Sud se rejoignent globalement. Les premiers témoignages concernent les ateliers de travail régionaux organisés pour assister les partenaires du Sud dans les phases d'élaboration des termes de référence et de mise en œuvre de l'évaluation. Ces ateliers de travail ont été jugés utiles et constructifs par les acteurs du Sud. Amacodou Ndiaye, directeur d'un cabinet sénégalais spécialisé dans le conseil en management et en évaluation déclare : « *le premier atelier a permis d'intégrer tous les pays de la région Afrique de l'Ouest dans l'élaboration des termes de référence, nous avons ainsi pu faire valoir les aspects que nous voulions voir traiter dans l'évaluation* ». Justine Odjoube, haut-fonctionnaire béninoise, estime que ces ateliers ont permis « *d'avoir une vision claire et commune de l'approche évaluative à suivre, ainsi qu'un cadre évaluatif unifié* ».

Un autre atout de l'évaluation régulièrement souligné est « *la création d'un 'groupe national de référence' chargé de coordonner l'évaluation et de contrôler la qualité du rapport final* ». Justine Odjoube y voit l'un des principaux facteurs d'amélioration par rapport aux évaluations partenariales auxquelles elle avait précédemment participé : « *pour l'évaluation de la Déclaration de Paris, les choses se sont améliorées (par rapport aux autres évaluations*

[819] Témoignages recueillis lors de l'organisation et du déroulement d'une conférence intitulée : *l'évaluation de l'aide au développement, quelles perspectives.* Pour le Bénin : Justine A Odjoube, Observatoire du Changement Social Bénin, ministère d'État chargé de la prospective, du développement et de l'Évaluation de l'Action Publique ; Damien Mededji, chercheur en économie, membre de l'équipe d'évaluation. Pour le Sénégal, Mayacine Camara, unité de coordination et de suivi de la politique économique, ministère de l'économie et des finances ; Amacodou Ndiaye, consultant, chef de l'équipe d'évaluation. Les évaluateurs du Sud ayant été associés à l'évaluation partenariale de la Déclaration de Paris ont donc des profils divers. Beaucoup sont des fonctionnaires qui sont chargés du suivi-évaluation des programmes économiques et sociaux de leur gouvernement, d'autres sont des chercheurs travaillant sur l'évaluation des politiques publiques ou des consultants indépendants connus des organismes d'aide, de par des collaborations passées sur des évaluations de programmes d'aide au développement. Ce sont des personnes expérimentées qui pratiquent l'évaluation depuis de nombreuses années.

partenariales) en ce sens que le groupe national de référence mis en place était composé d'experts béninois très qualifiés sur les domaines couverts par l'évaluation, ce qui a permis l'assurance qualité du rapport avant sa transmission au secrétariat de l'évaluation. La partie nationale a suivi de bout en bout le travail du consultant, lui a fourni les appuis nécessaires et a organisé l'atelier de validation du rapport »[820]. Damien Mededji, docteur en économie, estime que « *l'engagement des décisionnaires politiques dans l'évaluation a été un facteur fondamental de sa bonne mise en œuvre. Leur implication a notamment permis de faciliter la communication avec les autorités locales et l'organisation des ateliers de travail. Elle a également été une condition primordiale de l'utilisation des résultats et des recommandations de l'évaluation* ». Nous voyons que les évaluateurs du Sud reprennent dans leurs discours les priorités et le vocabulaire des évaluateurs du Nord : contrôle qualité des évaluations, utilisation des recommandations, cadre évaluatif, etc. Ainsi, un effet d'apprentissage semble bien avoir eu lieu, ou au moins un effet de mimétisme. Toutefois, les experts impliqués dans l'évaluation étaient pour la plupart les référents locaux habituels des donateurs en matière d'évaluation. Ils étaient donc déjà rompus aux techniques d'évaluation. En conséquence, il est difficile de juger du réel effet d'apprentissage permis par cette évaluation qui aurait été plus visible si des novices en évaluation avaient été intégrés au dispositif. Cependant, cela aurait menacé la qualité de l'évaluation qui s'est déroulée dans des contraintes de temps et méthodologiques importantes.

En effet, les évaluateurs les moins rompus aux évaluations partenariales témoignent avoir eu des difficultés à appréhender les temporalités des différentes phases de l'évaluation. Le processus de recrutement des consultants a été perçu comme trop long et les procédures de passation des marchés trop complexes, ce qui a entraîné un raccourcissement de la phase de mise en œuvre de l'évaluation et la difficulté de respecter les délais de rendus des livrables. Une autre difficulté, identifiée par Justine Odjoube, a été de « *mobiliser sur le long terme les fonctionnaires locaux qui trouvaient l'investissement demandé par l'évaluation trop lourd, d'autant plus qu'un certain nombre d'entre eux avaient déjà été sollicités pour les enquêtes de suivi* ». Leur implication était donc assez irrégulière, ce qui a pu résulter en une collecte des données qualitatives moins représentatives que ce qui était recherché au départ et en un affaiblissement du renforcement des capacités en évaluation. Mayacine Camara, membre du ministère sénégalais de l'économie constate « *une perte significative de la mémoire institutionnelle locale en raison de la rotation rapide des agents administratifs dans certaines structures sectorielles comme la santé ou l'éducation* ». Pour cette raison, ou pour des raisons

[820] Entretien avec la directrice de l'Observatoire du changement social du Bénin, 2 décembre 2013.

plus politiques, l'accès à certaines parties prenantes des programmes évalués a été limité, voire impossible. Ainsi, si l'engagement politique de haut niveau est apparu meilleur que dans les autres évaluations, sûrement en raison du retentissement de l'évaluation sur la scène internationale, l'implication des fonctionnaires des administrations est restée faible.

Concernant la méthodologie de l'évaluation, certains évaluateurs du Sud déclarent avoir été dépassés par sa complexité, ainsi que par le grand nombre d'indicateurs à renseigner. Ils avouent que toutes les matrices n'ont pas été remplies. Le blocage ressenti face à la méthodologie a été renforcé, d'après leurs témoignages, par « *une relation insuffisante avec l'équipe de coordination de l'évaluation au niveau international* », notamment en raison de la « *barrière linguistique* » et de « *difficultés organisationnelles* ». Le fait que l'évaluation soit mise en œuvre simultanément dans plusieurs dizaines de pays, sous l'égide d'une coordination internationale, était un défi d'envergure qui a entraîné une lourdeur bureaucratique considérable, dont ont souffert aussi bien les évaluateurs sur le terrain que les coordinateurs. Ce dispositif de coordination était toutefois indispensable pour assurer la cohérence et l'harmonisation de l'évaluation globale, afin qu'elle ne soit pas qu'une compilation de « rapports pays », mais bien une réflexion transversale sur les résultats de la Déclaration de Paris.

Cette analyse de la mise en œuvre de l'évaluation de la Déclaration de Paris montre que son caractère participatif reste imparfait et reproduit en partie les manquements habituels des évaluations partenariales. Si la participation des acteurs du Sud a été plus poussée que dans la plupart des autres cas d'évaluations partenariales, les individus engagés en tant qu'évaluateurs étaient souvent déjà très habitués à ce type de collaboration avec des évaluateurs et des donateurs occidentaux. Les moins expérimentées des parties prenantes du Sud sont restées à l'écart, dépassées par l'ampleur et la méthodologie de l'évaluation. Toutefois, cela a été une opportunité pour les officiels locaux de participer à une évaluation internationale et les évaluateurs du Sud ont pu tisser ou renforcer des liens professionnels avec des évaluateurs du Nord.

Cette première expérience d'évaluation participative au niveau mondial, incarnation du principe de responsabilité mutuelle promu par l'agenda sur l'efficacité de l'aide, dévoile bien le paradoxe actuel de l'évaluation, entre une volonté d'utiliser les méthodologies les plus complexes et avancées pour répondre aux exigences de rigueur et de redevabilité, et le devoir d'adopter une démarche participative plus qualitative qui soit le point de départ pour un dialogue ouvert sur l'aide au développement et un apprentissage mutuel. Ces deux approches

sont difficiles à combiner, dans la mesure où l'écart de capacités en évaluation entre les États du Nord et du Sud, mais aussi au sein de ces deux groupes d'États, reste important.

Toutefois, le jugement porté sur cet exercice ne doit pas être trop sévère, car il s'agit d'une expérimentation et qu'elle a permis aux acteurs du Nord de prendre conscience des ambiguïtés de leur approche évaluative. Cette initiative a pris place alors que la plupart des donateurs bilatéraux restent peu familiers avec l'évaluation partenariale. De plus, l'échelle internationale ne nous apparaît pas être la plus adéquate pour entraîner des effets d'apprentissage. Par ailleurs, cette expérimentation a donné naissance à un travail collectif au sein du réseau d'évaluateurs CAD-OCDE sur les opportunités ouvertes par les évaluations partenariales en termes de renforcement des capacités. Elle a incité les donateurs à renforcer la pratique de l'évaluation partenariale à des fins d'apprentissage mutuel. Selon Justine Odjoube, « *depuis l'évaluation de la Déclaration de Paris, de réels efforts sont fournis par les donateurs pour faire exécuter des évaluations partenariales incluant de plus en plus les acteurs et les responsabilisant, bien que les effets d'apprentissage ne soient pas toujours visibles* »[821]. L'enjeu est de déterminer sur quels aspects les évaluations partenariales sont améliorables pour permettre un meilleur renforcement des capacités et, au contraire, quels sont les changements que les donateurs ne pourront pas accepter compte tenu de leurs propres contraintes. Le processus de transformation du système évaluatif international, qui permettrait de rendre les évaluations partenariales effectives en matière de renforcement des capacités, en est encore à ses balbutiements.

C. Des systèmes d'évaluation désincarnés

Après avoir étudié les approches des bailleurs de fonds en matière de renforcement des capacités en évaluation, nous allons nous intéresser à la manière dont les pays bénéficiaires de l'aide s'approprient et mettent en place dans leurs politiques publiques les standards d'évaluation transférés par les donateurs.

La grande majorité des pays du Sud semblent s'engager dans des processus d'institutionnalisation de l'évaluation des politiques publiques. Dans de nombreux pays, nous observons l'apparition d'évaluateurs, de structures gouvernementales, de cadres législatifs et

[821] Entretien avec la Directrice de l'Observatoire du changement social du Bénin, 2 décembre 2013.

d'organisations professionnelles de « suivi-évaluation ». Dans leur littérature grise, les organismes de développement sont nombreux à présenter des cas de pays du Sud engagés dans la mise en place de systèmes nationaux d'évaluation et à analyser leurs réussites et leurs limites[822]. Toutefois, il existe peu de travaux essayant de comprendre les motivations qui poussent les gouvernements à mettre en place de tels systèmes et la manière dont ils conçoivent l'évaluation à l'aune de leur culture politique nationale.

- **Une poussée d'institutionnalisation de la fonction d'évaluation au Sud**

Les systèmes de suivi-évaluation qui émergent dans les pays bénéficiaires de l'aide sont souvent présentés par les organismes d'aide au développement, comme provenant de la volonté des gouvernements d'améliorer l'efficacité de leurs politiques publiques et d'une pression interne de la société civile en demande d'une meilleure transparence de la sphère politique[823]. Nous défendons la thèse selon laquelle cette vision est faussée dans la plupart des cas. Nous allons démontrer que les systèmes de suivi-évaluation des pays du Sud ont, avant tout, été créés pour satisfaire une demande des donateurs et que l'appropriation locale et l'utilisation de ces systèmes sont en réalité très faibles.

Pour mener notre analyse, nous nous basons sur les récents travaux de l'organisme international *Regional Center for Learning on Evaluation Results* (CLEAR). CLEAR est une initiative internationale de donateurs et de pays du Sud créée pour « renforcer les capacités et systèmes de suivi-évaluation et de gestion par la performance dans les États du Sud, afin de soutenir leurs efforts vers des politiques publiques axées sur les résultats »[824]. Nous retrouvons dans cette description le vocabulaire normé des organismes internationaux de développement : « systèmes de suivi-évaluation », « gestion par la performance », « axées sur les résultats ». L'approche développée par CLEAR consiste à mettre en place un renforcement des capacités davantage centré au Sud, principalement fondé sur le partage d'expériences et l'émulation Sud-Sud. CLEAR souhaite sortir du schéma Nord-Sud de transferts des compétences. L'initiative compte cinq pôles régionaux - Afrique anglophone, Afrique francophone, Amérique latine, Asie du Sud et Asie de l'Est - gérés par des experts régionaux en évaluation. Toutefois, le

[822] Voir par exemple, BOYLE, Richard, LEMAIRE, Donald, 1999, *op. cit ;* World Bank OED, *Building Country Capacity for Monitoring and Evaluation in the Public Sector: Selected Lessons of International Experience*, ECD Working Paper Serie, n°13, 2005.
[823] Entretien avec le directeur d'un cabinet d'évaluation au Sénégal, 3 avril 2014 ; entretien avec la coresponsable du réseau d'évaluateurs IOCE, 22 janvier 2014.
[824] Site internet de CLEAR : http://www.theclearinitiative.org/, consulté le 30 décembre 2013.

renouveau qu'apporte CLEAR est à relativiser, son objectif restant clairement d'instaurer « la gestion par les résultats » dans les États du Sud, concept fortement promu par les organisations d'aide au développement. Si la localisation de CLEAR est centrée au Sud, ses objectifs reflètent donc une influence forte des donateurs. CLEAR doit d'ailleurs son existence aux financements de nombreux bailleurs de fonds : la Banque mondiale, des banques régionales de développement, le DFID, la SIDA, les ministères des affaires étrangères suisse et australien et la fondation Rockefeller.

Un des premiers travaux de CLEAR a été d'élaborer des « fiches pays »[825] faisant le bilan des besoins, des compétences et des structures en évaluation, ainsi que de l'utilisation des évaluations dans quinze pays d'Afrique subsaharienne, à partir d'observations empiriques et d'entretiens auprès des officiels et de la société civile[826]. Ce travail a servi de base à notre recherche. Toutes les enquêtes par pays ont suivi la même méthodologie qui a été développée par un consultant américain, Carlisle Levine, docteur en relations internationales ayant 23 ans d'expérience en tant qu'évaluateur auprès d'organismes de développement, tels qu'USAID et CARE international. Malgré une approche Sud-Sud, CLEAR a donc préféré amorcer son travail analytique en s'appuyant sur un expert confirmé du Nord. Un consultant local a ensuite été sélectionné dans chaque pays pour mener l'étude de terrain. Les enquêtes portent sur l'ensemble des parties prenantes des systèmes de suivi-évaluation : les fonctionnaires aux niveaux central et sectoriel de l'administration, les évaluateurs indépendants, les représentants de la société civile et les universitaires. Les résultats des enquêtes ont ensuite été restitués aux parties prenantes nationales pour validation. Quatre aspects de la culture de l'évaluation ont été étudiés dans chaque pays :

- l'engagement de la classe politique à développer des capacités de suivi-évaluation ;
- l'existence d'une offre compétente de suivi-évaluation ;
- l'existence d'une demande et d'une utilisation des informations de suivi-évaluation ;
- l'existence d'un environnement favorable (réformes de la gouvernance, renforcement de la société civile, recherche académique de haut niveau et indépendante du politique, etc.).

[825] CLEAR-CESAG, *An Assessment of Demand for Monitoring, Evaluation, and Performance Management Information and Services*, Prepared for the CLEAR Francophone Africa Center at the Centre africain d'études supérieures en gestion, 2013; CLEAR, *Demand and Supply: Monitoring, Evaluation, and Performance Management Information and Service in Anglophone Sub-Saharan Africa*, 2013.

[826] Afrique du Sud, Bénin, Botswana, Ghana, Kenya, Malawi, Mauritanie, Namibie, Rwanda, Sénégal, Swaziland, Tanzanie, Uganda, Zambie, Zimbabwe.

L'objectif final de ce parangonnage de l'état de l'évaluation des politiques publiques dans les pays d'Afrique subsaharienne est d'élaborer des stratégies de renforcement des capacités en évaluation adaptées à chaque contexte. Cette réflexion participative devrait permettre aux gouvernements d'identifier leurs besoins et les forces et faiblesses de leurs systèmes actuels d'évaluation. L'objectif de CLEAR est également de démontrer aux donateurs qu'il existe bien des compétences en évaluation dans les États du Sud et donc qu'il est possible de mettre en œuvre des évaluations partenariales sur des bases plus équitables[827]. L'objectif politique implicite est donc de rendre plus égalitaire la relation entre les évaluateurs du Sud et du Nord et plus largement entre les donateurs et les bénéficiaires dans le système de l'aide. Si CLEAR a fait, à partir de ces enquêtes, une bonne description des systèmes d'évaluation dans les pays africains, il ne va pas jusqu'à offrir une analyse des causes et conséquences politiques des constats réalisés. C'est ce que nous nous proposons de faire ici.

À partir des « fiches pays » réalisées par CLEAR et de recherches complémentaires, nous avons étudié les modalités des systèmes d'évaluation mis en place par les États d'Afrique subsaharienne, observé leur degré d'inscription dans les politiques publiques nationales et estimé leur durabilité. En nous situant dans le courant néo-institutionnaliste de la sociologie des organisations, nous partons du postulat que l'adoption de systèmes d'évaluation nationaux par les gouvernements du Sud est moins une question « de stratégies scientifiquement élaborées, par lesquelles les importateurs (du transfert de bonnes pratiques) cherchent à accroître leur efficacité et leur efficience, que l'expression d'un besoin de légitimation »[828]. En créant des systèmes d'évaluation, les gouvernements du Sud se conforment aux standards internationaux de gestion des politiques publiques. Ils cherchent ainsi à « susciter la confiance et le soutien des acteurs dont ils dépendent pour poursuivre leurs activités »[829], ici les bailleurs de fonds qui financent une partie de leurs politiques publiques. Cette lecture implique que l'objectif des gouvernements, lorsqu'ils établissent des systèmes de suivi-évaluation, est, en premier lieu, de convaincre les donateurs de leur fiabilité.

Dans les « fiches pays » réalisées par les évaluateurs de CLEAR, il est frappant de constater que la demande en évaluation exprimée par les gouvernements d'Afrique subsaharienne semble toujours très liée aux exigences des bailleurs de fonds internationaux, ce

[827] Hadji Gueye, Docteur en économie, Coordonnateur du Centre CLEAR de l'Afrique francophone, allocution lors d'un séminaire de la société française d'évaluation intitulée « quelles perspectives pour l'évaluation de l'aide au développement » organisée le 13 février 2013 à Paris.
[828] MEYER, John. W, ROWAN, Brian, "Institutionalized Organizational Analysis: Formal Structure as Myth and Ceremony", in POWELL, Walter. W, Di MAGGIO, Paul. J, *The New Institutionalism in Organizational Analysis*, Chicago, University Press of Chicago Press, 1991, 486p, p. 43.
[829] *Ibid*.

qui confirme la pertinence de notre lecture de ce phénomène. Au Sénégal, par exemple, nous observons qu'il y a une faible demande nationale d'évaluation en dehors des projets financés par les organismes internationaux. En Mauritanie, nous remarquons que le moteur du suivi-évaluation réside dans les exigences des donateurs[830]. L'influence normative des donateurs se manifeste aussi par le fait que les engagements sur la mise en place de systèmes nationaux de suivi-évaluation ont été faits dans le cadre des « stratégies de réduction de la pauvreté » et des « documents de planification nationale », documents politiques dont le contenu conditionne les allocations d'aide au développement. Les pays bénéficiaires, avertis des exigences des donateurs, affichent donc dans ces documents leur volonté de renforcer leurs capacités en suivi-évaluation.

La construction de systèmes d'évaluation s'opère souvent dans un ensemble plus large de réformes de la gouvernance avec l'adoption de lois et de réglementations allant vers plus de transparence et de redevabilité des politiques publiques, prérequis pour créer un environnement favorable au développement d'une culture de l'évaluation. Par exemple, au Rwanda, le gouvernement a lancé, en 2001, la *National Information and Communication Infrastructure Policy*, qui vise à développer l'utilisation des technologies informatiques dans le service public et le secteur privé[831]. En Ouganda, le gouvernement a établi, en 2003, le *Public Finance Accountability Act*, qui requiert du ministère des finances qu'il rende compte au Parlement sur les dépenses publiques de l'État[832]. Ces documents présentent des réformes en termes de gestion axée sur les résultats, de transparence et de suivi des dépenses publiques. D'autres pays, le Ghana, le Kenya, ou encore l'Afrique du Sud, sont allés plus loin en inscrivant le principe du suivi-évaluation des politiques publiques dans leurs constitutions. Certains de ces pays ont élaboré (l'Afrique du Sud) ou sont en train d'élaborer (le Kenya et l'Ouganda) des politiques nationales d'évaluation[833]. Selon les hauts fonctionnaires de ces pays, « *élaborer une politique nationale d'évaluation, ou en faire un élément constitutionnel, permet d'établir des standards de qualité et d'éthique pour l'évaluation, d'harmoniser les pratiques et de clairement définir les responsabilités des différents acteurs publics en matière de suivi-évaluation* »[834]. Si en Afrique du Sud et en Ouganda, la demande de la société civile semble avoir été un vecteur important pour l'intégration de l'évaluation dans la constitution, ce n'est pas le cas au Kenya,

[830] CLEAR-CESAG, 2013, fiches pays Mauritanie et Sénéral, *op. cit.*
[831] CLEAR, 2013, fiche pays Kenya, *op. cit.*
[832] CLEAR, fiche pays Ouganda, 2013, *op. cit.*
[833] *Ibid*.
[834] Hadji Gueye, docteur en économie, coordonnateur du centre CLEAR de l'Afrique francophone, allocution lors d'un séminaire de la société française d'évaluation intitulée « quelles perspectives pour l'évaluation de l'aide au développement » organisée le 13 février 2013 à Paris.

où cette décision semble davantage relever de l'influence de la Banque mondiale. Dans ces deux premiers pays, nous constatons d'ailleurs que la demande d'évaluation des gouvernements tend à s'élargir de manière significative à d'autres programmes de politiques publiques que ceux financés par les donateurs, ce qui est moins le cas au Kenya.

L'adoption de ces cadres législatifs va de pair avec la création de structures chargées de mettre en œuvre le suivi-évaluation. Chacun des pays étudiés a mis en place une ou plusieurs structures gouvernementales dédiées au suivi-évaluation. Dans la plupart des pays, nous constatons que les entités responsables du suivi-évaluation au niveau gouvernemental sont très nombreuses. Nous pouvons prendre l'exemple de l'Afrique du Sud, où la fonction d'évaluation est partagée, au sein du gouvernement, entre : le département de la performance, du suivi et de l'évaluation au sein de la Présidence, le Trésor, le département du service public et de l'administration, l'auditeur général, le département de la gouvernance coopérative, le bureau du Premier ministre, la commission du service public et la commission de la planification nationale[835].

À cela s'ajoutent des structures sectorielles d'évaluation souvent créées pour répondre aux besoins des donateurs de suivi-évaluation de leurs programmes d'aide. Ces structures fonctionnent généralement mieux que celles de niveau gouvernemental, spécifiquement dans les secteurs de la santé et de l'éducation, comme c'est le cas au Botswana, en Namibie, au Rwanda, au Sénégal, ou encore au Bénin. En effet, ce sont des secteurs d'activité sur lesquels les donateurs interviennent massivement et évaluent beaucoup leurs interventions[836].

Enfin, nous retrouvons des structures *ad hoc* de suivi-évaluation au niveau des projets financés par des bailleurs de fonds internationaux.

Ces trois « strates » de suivi-évaluation – gouvernementale, sectorielle et projet - se sont développées sans harmonisation, ce qui entraîne des problèmes de coordination des activités de suivi-évaluation au niveau national. Leur coordination constitue un défi pour les gouvernements, afin de pouvoir capitaliser les connaissances apportées par les travaux d'évaluation et harmoniser les compétences des évaluateurs et la qualité de leurs travaux.

La fragmentation d'un instrument de politique publique en un grand nombre d'acteurs ou de réseaux est souvent identifiée dans les études sur les *policy transfers*[837]. Dans ces travaux, il

[835] CLEAR, 2013, fiche pays Afrique du Sud, *op. cit.*
[836] *Ibid*.
[837] DEZALAY, Yves, GARTH, Bryant. G, *Global Prescription. The Production, Exportation, and Importation of a New Legal Orthodoxy*, Ann Arbor, University of Michigan Press, 2002, 369p, p. 28.

est constaté, comme c'est le cas dans notre recherche, que les acteurs de chaque strate disposent d'une large autonomie pour adapter à leurs besoins, leurs compétences et leurs contraintes les instruments de politiques publiques importés. En conséquence, au sein d'un même État, des pratiques divergentes peuvent émerger entre ces strates en fonction de l'intérêt des acteurs pour l'instrument importé, de leur compréhension de cet instrument, de leur budget, etc.[838].

Cette analyse nous permet de comprendre que les efforts des États d'Afrique subsaharienne se sont principalement concentrés sur l'institutionnalisation de la fonction d'évaluation dans l'administration. Ils ont donc reproduit l'approche *top down* du renforcement des capacités en évaluation des organismes d'aide au développement, selon laquelle il faut d'abord créer un cadre (lois et institutions) et une offre d'évaluation qui vont ensuite stimuler une demande en évaluation. C'est une caractéristique courante des processus de transferts de politiques publiques, relevée par Eleanor Westney : « une cause de dénaturation (du modèle importé) réside dans la focalisation des entrepreneurs de transfert sur certains aspects du modèle, alors que d'autres sont négligés. Dans la mesure où la transposition est ordinairement planifiée par le sommet de l'organisation et par ses tutelles, une attention plus grande est portée aux dispositifs et aux procédures officiels de contrôle et de pilotage, aux règles de gestion, à la structure hiérarchique et aux modalités d'acquisition des ressources, qu'au déroulement des activités concrètes à la base des organisations »[839]. Ainsi, la manière dont la politique ou l'instrument va pouvoir être « opérationnalisé » dans le contexte d'importation est négligée, ce qui rend difficile sa mise en pratique par les acteurs nationaux, malgré une institutionnalisation qui peut être très poussée, comme c'est le cas en Afrique du Sud. Les individus désignés comme étant chargés de l'évaluation ne sont pas guidés par une demande claire des dirigeants politiques. Il n'y a pas de dispositif pour favoriser les interactions entre ces évaluateurs et les administrations. Ainsi, on observe le même type de blocages que dans les organismes d'aide occidentaux.

Il convient alors de se demander si l'institutionnalisation du suivi-évaluation dans les pays d'Afrique subsaharienne, qui donne l'impression d'une amélioration de la culture de l'évaluation, n'est pas en réalité une coquille vide.

- **Des systèmes d'évaluation mis en place pour les donateurs**

[838] *Ibid*.
[839] WESTNEY, Eleanor, *Imitation and Innovation: the Transfer of Western Organizational Patterns to Meiji Japan*, Cambridge, Harvard University Press, 1987, 256p, p. 7.

Premièrement, nous nous sommes questionnée quant à la fonctionnalité réelle des lois et des structures fondatrices de l'activité d'évaluation au Sud. Dans l'ensemble des pays, la mise en pratique du suivi-évaluation apparaît faible. Les capacités en suivi-évaluation sont rares et éparpillées dans de multiples entités et le personnel en charge de l'évaluation est jugé insuffisant[840]. Compte tenu de ce manque de capacités et de ressources humaines dans les institutions publiques, les gouvernements externalisent souvent les activités de suivi-évaluation à des consultants ou à des instituts de recherche. Les ressources budgétaires allouées au suivi-évaluation sont jugées insuffisantes et nous remarquons qu'elles sont, en large majorité, allouées au suivi, au détriment de l'évaluation. Au Ghana notamment, sur un budget total de suivi-évaluation d'approximativement onze millions de dollars[841] en 2010, 70 % revenait au suivi[842]. En effet, si les structures nationales de suivi-évaluation mettent régulièrement en place des activités de suivi, les évaluations sont encore très majoritairement menées par les donateurs. Ainsi, les officiels kényans interrogés pour l'étude de CLEAR relatent que « le système de suivi-évaluation du Kenya se concentre plus sur le suivi que sur l'évaluation. Les évaluations sont surtout conduites par les donateurs »[843]. En Afrique du Sud, les témoignages sont similaires : « l'évaluation est encore largement gérée par les donateurs. Les efforts du gouvernement restent concentrés sur le suivi »[844].

Nous expliquons cette situation par le fait que l'évaluation est historiquement une obligation imposée par les donateurs à des fins de contrôle sur la manière dont les gouvernements dépensent les fonds d'aide et conçoivent leurs politiques publiques. Les systèmes d'évaluation ont été mis en place sans qu'un travail de sensibilisation et d'information sur la fonction et l'utilité de l'évaluation n'ait été mené dans les administrations. Selon Justine Odjoube, haut fonctionnaire béninoise, « *l'évaluation au Bénin a été conduite historiquement à la demande des donateurs et exécutée, la plupart du temps, par des consultants externes étrangers, avec une participation variable de consultants nationaux. Le recours à l'évaluation n'est pas régulier et l'inexistence de budget pour conduire les évaluations en est la principale raison. Les évaluations sont réalisées en général, soit en réponse aux exigences des bailleurs de fonds, soit par obligation réglementaire ou légale et, dans une moindre mesure, sur initiative propre des administrations* »[845]. Il n'y a donc pas réellement de demande nationale d'évaluation. Les gouvernements se sont dépêchés de mettre en place les éléments « visibles »

[840] CLEAR-CESAG, 2013, *op. cit.* CLEAR, 2013, *op. cit.*
[841] Principalement composé d'aide budgétaire internationale.
[842] CLEAR, fiche pays Ghana, 2013, *op. cit.*
[843] CLEAR, fiche pays Kenya, 2013, *op. cit.*
[844] CLEAR, fiche pays Afrique du Sud, 2013, *op. cit.*
[845] Entretien avec la Directrice de l'Observatoire du changement social du Bénin, 2 décembre 2013.

(institutions, régulations) des systèmes nationaux d'évaluation pour afficher leur conformité aux bonnes pratiques internationales de gouvernance, mais ont attaché une importance moindre à l'opérationnalisation de ces systèmes, à leur adaptation au contexte national et à leur appropriation par les acteurs des services publics. Schéma que l'on retrouvait dans les organismes d'aide au développement du Nord dans les années 1980.

En conséquence, nous observons que les agents de la fonction publique ont une vision partielle et plutôt négative de l'évaluation, qui est vue comme un instrument de contrôle ne pouvant que leur nuire, alors que ses fonctions d'apprentissage et de démocratisation sont ignorées. En Mauritanie, il ressort que l'utilité de l'évaluation est mal comprise, en raison d'un accent trop fort mis par le gouvernement sur le contrôle qui empêche les acteurs locaux de percevoir les effets d'apprentissage que peut entraîner l'évaluation[846]. Au Kenya, il apparaît également que l'évaluation est plus vue comme un instrument de contrôle que comme un moyen d'améliorer l'efficience et l'efficacité des politiques publiques[847].

Cette méconnaissance de la part des opérateurs et des facilitateurs nationaux de transfert[848] de la bonne pratique qu'ils importent n'est pas spécifique à l'Afrique subsaharienne, ni au suivi-évaluation. Selon Harold Wolman, « la plupart du temps, l'importateur n'a qu'une compréhension très lacunaire des interactions et des interdépendances systémiques qui font que la solution est efficace dans l'environnement d'où elle est extraite, souvent parce qu'il n'a pas su trouver ou sélectionner les données pertinentes nécessaires, ou s'il en dispose, parce qu'il lui manque des compétences requises, ou le temps suffisant, pour les interpréter. Les conditions qui doivent être réunies pour que la transposition ait de bonnes chances de produire les résultats attendus ne sont pas recensées de manière systématique et rigoureuse. Peu d'efforts sont faits pour identifier les incompatibilités liées au contexte d'accueil et pour estimer les coûts d'adaptation de celui-ci à l'introduction des éléments importés »[849]. Dans le cas de notre objet d'étude, nous expliquons cette tendance par le fait que les motivations relevant de considérations de réputation internationale et de pressions normatives externes aux États l'emportent sur les motivations d'amélioration des politiques publiques dans la volonté des gouvernements à importer des bonnes pratiques de suivi-évaluation.

Un autre indicateur de l'ancrage d'une culture de l'évaluation dans un État est l'utilisation qui est faite du suivi-évaluation. Il est difficile de mesurer le degré d'utilisation de l'évaluation,

[846] CLEAR-CESAG, fiche pays Mauritanie, 2013, *op. cit.*
[847] CLEAR, fiche pays Kenya, 2013, *op. cit.*
[848] Voir typologie de Wolman ci-dessus.
[849] WOLMAN, Harold, 2002, *art. cit.* p. 498.

car elle est souvent indirecte et se réalise sur le long terme. Notre analyse s'est donc essentiellement fondée sur les témoignages des officiels interrogés lors des études menées par CLEAR dans chaque pays.

À partir de ces témoignages, nous constatons, qu'hormis au Malawi et en Zambie, aucune des personnes interrogées ne témoigne d'une utilisation régulière des résultats du suivi-évaluation[850]. Lorsqu'il y a utilisation, elle se limite à l'appui à la détermination du budget des administrations. Ce constat nous amène à penser que la culture de l'évaluation est faible, y compris dans les pays où l'institutionnalisation des systèmes de suivi-évaluation est la plus aboutie. Même en Afrique du Sud, État considéré par les experts en évaluation comme un « bon élève », il ressort de notre étude que la culture de l'évaluation est encore peu développée, qu'il y a un manque d'appropriation du système de suivi-évaluation par ses opérateurs, qu'il règne une forme de « conformité malveillante »[851] aux standards de suivi-évaluation et que les administrations développent peu d'intérêt pour l'utilisation des informations fournies par les évaluations[852]. Nous en concluons que les systèmes de suivi-évaluation nationaux des pays africains sont désincarnés et peu appropriés par les acteurs locaux.

Cette conclusion va dans le sens des théories élaborées par les tenants du courant néo-institutionnaliste de la sociologie politique à propos des *policy transfers* qui expliquent que « le transfert est essentiellement conditionné par la volonté du gouvernement importateur de ne pas apparaître comme « retardataire », quitte à ce que la greffe soit opérée de manière purement rituelle, sans souci de cohérence avec ses propres activités. Ce souci de conformité aux modèles dominants est particulièrement prégnant dans les organisations telles que les administrations publiques, dont les performances ne peuvent être évaluées de façon claire et univoque et où le respect d'obligations de moyens constitue le principal critère de « bon fonctionnement »[853]. En effet, la mise en place de systèmes d'évaluation nationaux pour les pays bénéficiaires de l'aide au développement a une forte portée symbolique. Ils envoient ainsi un signal de bonne gouvernance et de confiance aux bailleurs de fonds. Elizabeth Heger Boyle et John Meyer ont théorisé cette fonction symbolique du phénomène d'alignement d'acteurs sur des standards internationaux : « il constitue un signe extérieur par lequel un État manifeste sa volonté d'appartenance à une communauté internationale, quand bien même les autorités nationales

[850] CLEAR, fiches pays Malawi et Zambie, 2013, *op. cit.*
[851] CLEAR, fiche pays Afrique du Sud, 2013, *op. cit.*
[852] Entretien avec une évaluatrice de la Banque mondiale, 24 octobre 2013.
[853] POWELL, Walter, W, Di MAGGIO, Paul. J, *The New Institutionalism in Organizational Analysis*, Chicago, University of Chicago Press, 1991, 486p.

importatrices n'ont pas la volonté de mettre effectivement en œuvre les normes transférées et ferment les yeux sur leur transgression »[854].

Selon nous, si la « greffe » n'a pas pris dans la grande majorité des cas d'études abordés, c'est parce que le lien n'a pas été créé entre la culture nationale de politiques publiques et les modèles de systèmes de suivi-évaluation importés, fortement inspirés des bonnes pratiques des donateurs étrangers. Les résistances nationales, actives ou passives, n'ont pas été intégrées dans la stratégie d'implantation des systèmes d'évaluation et constituent aujourd'hui un blocage majeur à leur fonctionnement.

L'importation d'un système d'évaluation renvoie à une transformation de la culture politique générale de l'État et pas seulement de la technique de gouvernement. C'est pourquoi un des facteurs de la réussite du processus d'importation est l'implication de la société civile dans ce processus. La plupart des gouvernements se sont engagés à intégrer les OSC dans le renforcement des capacités en évaluation. Un signe apparemment positif en ce sens est la présence dans chaque pays étudié d'une ou de plusieurs sociétés professionnelles d'évaluation. Toutefois, certaines d'entre elles ne sont pas reconnues par le gouvernement. De plus, la plupart sont décrites comme inactives, à l'exception de l'Ouganda et de l'Afrique du Sud[855]. Généralement, la société civile reste coupée des réflexions et des activités gouvernementales en matière de suivi-évaluation. Cette situation est en partie due à la faible structuration de la société civile dans les pays étudiés, mais montre également que le suivi-évaluation n'est pas envisagé par les pouvoirs publics comme un outil de démocratisation et de dialogue politique.

Cependant, des OSC qui ne sont pas directement en lien avec les activités de suivi-évaluation peuvent contribuer au développement d'une culture de l'évaluation. Nous le constatons, par exemple, avec les organisations travaillant pour la lutte contre le sida qui conduisent des enquêtes permettant de renforcer les systèmes statistiques nationaux et des évaluations. Enfin, les universités et les instituts de recherche jouent un rôle dans la diffusion d'une culture de l'évaluation en fournissant des formations sur ce thème dans la majorité des cas d'étude.

Notre analyse nous a permis de comprendre les logiques qui animent les gouvernements des pays du Sud dans l'adoption de systèmes de suivi-évaluation. Les gouvernements d'Afrique subsaharienne apparaissent encore très dépendants des donateurs dans leurs dynamiques de

[854] HEGER BOYLE, Elizabeth, MEYER, John W, "Modern Law as a Secularized Global Model: Implications for the Sociology of Law", in DEZALAY, Yves, GARTH, Bryant. G, *Global Prescription. The production, Exportation, and Importation of a New Legal Orthodoxy*, Ann Arbor, University of Michigan Press, 2002, 368p.
[855] CLEAR, fiche pays Afrique du Sud et Ouganda, 2013, *op. cit.*

développement de tels systèmes. Les donateurs se satisfont de cette dépendance, car cela leur permet de garder un contrôle étroit sur la mise en œuvre des évaluations des projets et des programmes d'aide. De plus, les hauts fonctionnaires locaux ne perçoivent généralement pas les bénéfices qu'ils peuvent tirer du suivi-évaluation. L'existence politique même des opérateurs ou des facilitateurs locaux des systèmes de suivi-évaluation, les ressources qu'ils emploient et les modalités d'action qu'ils mettent en œuvre sont principalement le résultat de l'intervention des bailleurs de fonds internationaux. Selon Thierry Delpeuch, il s'agit « d'une stratégie couramment employée par les entrepreneurs de transferts disposant d'abondantes ressources, qui consiste à provoquer l'émergence d'acteurs nationaux, tant étatiques que non-étatiques, fortement disposés à soutenir l'importation des solutions que préconisent ces entrepreneurs, ainsi qu'à orienter l'activité de ces « agents de transfert » locaux par le biais de l'allocation de moyens, la promotion de conceptions et l'apprentissage d'expertise »[856]. Les systèmes de suivi-évaluation et leurs acteurs sont donc davantage les produits des bailleurs de fonds internationaux que d'une dynamique nationale de réformes des politiques publiques.

En conséquence, la durabilité de ces systèmes de suivi-évaluation, qui manquent d'une « incarnation » par les acteurs des administrations et de la société civile, est loin d'être assurée. Leur incorporation dans les politiques publiques nationales reflète l'approche des donateurs en matière de renforcement des capacités en évaluation avec un accent fort mis sur la redevabilité et la gestion axée sur les résultats et une focalisation sur l'institutionnalisation et le développement d'une offre en évaluation. La stratégie la plus commune des importateurs est d'adopter *a minima* les éléments les plus visibles de des bonnes pratiques d'évaluation, tout en essayant de conserver le fonctionnement habituel de leur système de politiques publiques. Le degré de conformité de l'importateur aux exigences de l'exportateur dépend alors de l'aptitude de ce dernier - souvent limitée - à déployer des moyens de suivi et de contrôle pour vérifier l'effectivité de la mise en œuvre des réformes.

Toutefois, il ne faut pas se situer dans une vision Nord/Sud restreinte qui donnerait l'image de transferts strictement imposés par des rapports de dépendance et de domination. Comme le souligne John Ikenberry, « la plupart du temps, les transferts internationaux ne relèvent pas de processus d'imposition unilatérale de politiques ou de pratiques, mais de la construction d'alliances entre l'exportateur et certains groupes d'intérêts et fractions des élites

[856] DELPEUCH, Thierry, 2008, *op. cit.* p. 46.

dirigeantes du système importateur »[857]. Comme nous l'avons vu, les importateurs disposent d'une marge de manœuvre large dans la manière dont ils appliquent les bonnes pratiques importées et savent en tirer des bénéfices : bonne réputation internationale, ressources d'aide supplémentaires, approbation de leur population, etc.

Dans les cas d'études abordés ici, il manque une phase d'adaptation et d'appropriation des structures et des cadres législatifs sur le suivi-évaluation pour qu'il soit possible de parler de diffusion d'une « culture » de l'évaluation. N'ont été transférés que les éléments concrets ou institutionnels de la culture de l'évaluation et non les « matériaux symboliques » de cette culture, tels que les idées, les systèmes de pensée ou les connaissances qui la fondent. Ce serait aux « opérateurs locaux du transfert », c'est-à-dire « les acteurs du pays importateurs qui ont acquis la maîtrise des discours et des savoirs employés par les « conseillers » étrangers »[858] de jouer ce rôle clé. Or, ces « opérateurs » ne sont pas structurés en une coalition dans les pays du Sud. Ils sont fragmentés au sein des administrations et sur le territoire en général, ne sont pas forcément spécialisés en évaluation et n'ont pas les moyens de diffuser leurs connaissances en l'absence d'un appui politique fort et continu. Justine Odjoube décrit bien cette situation au Bénin : « *il existe très peu de cabinets d'études spécialisés en évaluation au Bénin. Chaque fois que le besoin se fait sentir, la fonction publique fait appel à des compétences individuelles en évaluation. Mais, la plupart de ces compétences individuelles exercent rarement la fonction d'évaluateurs en tant que profession principale* »[859].

Nous identifions également un phénomène d'émulation entre les pays d'Afrique subsaharienne en matière d'adoption de systèmes d'évaluation. Une fois qu'un groupe de pays a adopté cette bonne pratique (la plupart du temps, les pays d'Afrique anglophone), les autres pays (d'Afrique francophone) ont suivi cette tendance, afin de ne pas apparaître comme des « retardataires » ou des « mauvais élèves » aux yeux des bailleurs de fonds internationaux, auprès desquels ils sont en concurrence pour toucher des financements. Comme le remarque Thierry Delpeuch, « dans beaucoup de cas, le simple fait qu'une politique soit mise en œuvre dans plusieurs autres pays suffit à justifier son importation »[860].

[857] IKENBERRY, John. G, "The International Spread of Privatization Policies: Inducements, Learning and "Policy Bandwagoning", in SULIEMAN, Ezra, WATERBURY, John. N, *The Political Economy of Public Sector Reform and Privatization*, Boulder, San Francisco, Oxford, Westview, 1990, 388p.
[858] EVANS, Mark, 2004, *op. cit.*
[859] Entretien avec la directrice de l'Observatoire du Changement social au Bénin, 2 décembre 2013.
[860] DELPEUCH, Thierry, 2008, *op. cit.*p. 33.

- **Conclusion de chapitre**

Dans ce chapitre, nous avons analysé la diffusion d'une culture de l'évaluation des organisations multilatérales et bilatérales vers les pays bénéficiaires de l'aide. Nous avons démontré que les vecteurs employés par les donateurs pour effectuer ce transfert (formations, assistance technique, évaluations partenariales) parviennent peu à stimuler une demande d'évaluation dans les pays en développement. Les donateurs restent dans une approche passive du transfert de bonnes pratiques politiques où les pays du Sud sont les « élèves » et les pays du Nord les « professeurs ». Les instruments actifs de renforcement des capacités en évaluation, sont encore très peu développés. Cependant, certains évaluateurs sont très impliqués dans l'amélioration de ce mode de transfert de compétences en évaluation qui pourrait donc devenir plus efficace dans les années à venir.

En conséquence, les pays du Sud adoptent généralement des réformes « de façade ». Ils mettent en place des structures et instaurent des lois sur le suivi-évaluation qui restent peu incarnées et peu opérationnelles. Cette situation convient aussi bien aux gouvernements du Sud, qui voient leurs pratiques des politiques publiques modifiées seulement à la marge, qu'aux donateurs, qui gardent le contrôle sur les activités de suivi-évaluation en restant les principaux acteurs de l'évaluation dans le Sud. Le poids des donateurs est donc fort dans l'institutionnalisation de la fonction d'évaluation dans les pays du Sud ce qui justifie de l'étudier dans le cadre d'une recherche sur l'aide au développement. Toutefois, au-delà d'un processus de transfert du Nord vers le Sud, nous avons observé un phénomène d'émulation entre pays du Sud qui provient des quelques pays dont les dirigeants se sont le plus engagés dans une transformation plus profonde de la culture politique nationale, comme en Afrique du Sud ou en Ouganda. Les sociétés professionnelles d'évaluation ancrées au Sud pourraient également être un vecteur de diffusion Sud-Sud d'une culture de l'évaluation des politiques publiques.

Dans le prochain chapitre, nous allons creuser cette piste et étudier les processus de diffusion horizontale d'une culture de l'évaluation au Sud qui se développent actuellement.

Chapitre 6 : Les réseaux d'évaluateurs : quelle influence normative et politique dans le régime de l'aide ?

Dans ce chapitre, nous nous intéresserons à la diffusion horizontale de la culture internationale de l'évaluation à travers l'action des sociétés professionnelles d'évaluateurs. Ces sociétés existent à différents niveaux : national, régional et international. Alors qu'elles n'étaient qu'une quinzaine au milieu des années 1990, elles sont aujourd'hui 114[861]. Ces organisations non gouvernementales, regroupant des individus et des organismes producteurs ou utilisateurs d'évaluations, ont de multiples dénominations à travers le monde : sociétés d'évaluation, réseaux d'évaluateurs, associations pour l'évaluation, communautés d'évaluateurs, ou encore forums sur l'évaluation. Dans cette étude, nous utiliserons indifféremment les termes de sociétés et de réseaux d'évaluation.

Face à cette multiplication des sociétés d'évaluation, nous nous sommes demandé quel est leur rôle dans la diffusion d'une culture de l'évaluation à travers le monde, notamment dans les pays du Sud. Nous voulions aussi tenter de cerner l'influence politique qu'elles peuvent avoir au niveau des pays, des régions ou international dans l'amélioration de la gouvernance de l'aide, mais aussi des politiques publiques des États. Dans un premier temps, nous étudierons le contexte international et les ambitions qui ont mené à leur création. Dans un deuxième temps, nous analyserons leurs modes de fonctionnement interne et les instruments qu'elles mobilisent pour agir dans le système international. Enfin, dans une troisième partie, nous essaierons d'estimer leur influence sur les politiques nationales et internationales. Nous conclurons ainsi sur les opportunités et les limites de ce mode de transfert horizontal, par les pairs, de bonnes pratiques d'évaluation.

A. Des réseaux instrumentalisés par les organisations multilatérales

Avant de commencer à étudier la genèse, le fonctionnement et les réalisations de ces réseaux d'évaluateurs dans l'espace international, il nous apparaît nécessaire de revenir sur la manière dont la théorie des relations internationales catégorise et définit ce type d'acteurs. La notion de réseau provient de l'anthropologie et de la sociologie urbaine et est de plus en plus

[861] Nous ne décomptons ici que les sociétés professionnelles d'évaluation « actives », c'est-à-dire qui ont des activités et/ou un site internet actif qui démontre que la structure n'est pas une organisation « fantôme ». Si l'on prend en compte l'ensemble des sociétés professionnelles d'évaluation actives et inactive, leur nombre total s'élève à plus de 150. Source : http://www.ioce.net/en/VOPEResources.php, consulté le 28 novembre 2014.

utilisée en relations internationales, notamment dans l'étude des interactions entre acteurs non gouvernementaux, qui se donnent la vocation d'apporter des voix alternatives à celles des États. Cette notion est souvent utilisée de manière dévoyée, de façon plus symbolique que précise et rigoureuse.

Ainsi, une de nos questions de recherche est de déterminer si ces sociétés d'évaluateurs, qui émergent partout dans le monde, forment réellement un réseau. Pour définir ce terme, nous référons tout d'abord à la notion de communautés épistémiques, définie par Peter Haas comme des « réseaux de professionnels ayant une expertise et une compétence reconnues dans un domaine particulier, qui peuvent faire valoir un savoir pertinent sur les politiques publiques du domaine en question. (…) Ce qui distingue les communautés épistémiques des autres groupes est la combinaison de croyances analytiques et normatives partagées, d'une base de connaissances consensuelles et d'intérêts communs »[862]. Un obstacle à l'identification des sociétés d'évaluateurs à ces communautés épistémiques est, comme nous allons le voir dans la suite de ce chapitre, que l'évaluation n'est pas considérée comme une profession à part entière dans la majorité des États. Toutefois, comme nous l'avons démontré dans les chapitres précédents, les évaluateurs partagent bien des croyances normatives, des connaissances consensuelles et des intérêts communs.

Margaret Keck et Kathryn Sikkink ont théorisé un type alternatif de réseaux internationaux : les *advocacy transnational networks*, qu'elles qualifient de « groupements d'acteurs motivés par des valeurs plutôt que par des intérêts matériels ou des normes professionnelles. Ces réseaux peuvent agir à la fois aux niveaux national et international. Les acteurs qui les composent travaillent sur un même enjeu et sont liés par des valeurs partagées, un discours commun et un échange important d'informations et de services »[863]. Le problème de cette définition est que même si les évaluateurs partagent un discours et des valeurs et même s'ils échangent des connaissances et leurs expériences au sein des réseaux qu'ils ont formés, ils sont liés par des normes et des intérêts professionnels.

Les sociétés d'évaluateurs seraient donc un hybride entre ces deux formes de réseaux internationaux. Elles sont constituées d'experts partageant des connaissances, des valeurs et des normes. Les évaluateurs se réunissent pour échanger et renforcer leurs capacités, ainsi que pour faire valoir leur expertise par des actions de plaidoyer et de professionnalisation. Toutefois, si

[862] HAAS, Peter, *Knowledge, Power and International Relations*, University of South Carolina Press, 1997, 404p, p. 13.
[863] KECK, Margaret, SIKKINK, Kathryn, *Activist beyond borders: Advocacy Networks in International Politics*, Cornwell University Press, 1998, 240p, p. 2.

une société d'évaluation peut former un réseau d'évaluateurs en elle-même, peut-on parler d'un réseau régional ou international qui regrouperait les différentes sociétés ? Les liens qu'elles entretiennent sont-ils étroits ou distendus ? Les évaluateurs du Nord et du Sud partagent-ils des intérêts et les sociétés d'évaluation sont-elles un bon vecteur d'échange de pratiques et de connaissances entre eux ?

- **Un contexte favorable à l'émergence d'associations d'évaluateurs**

Les premières sociétés nationales d'évaluation à être créées furent la Société canadienne d'évaluation en 1981, puis la Société américaine d'évaluation en 1986. Au niveau international, le CAD-OCDE mit en place, en 1986, son premier réseau d'évaluateurs, le *DAC Working Group on Evaluation*. Jusqu'au milieu des années 1990, le nombre de ces sociétés d'évaluateurs stagna, se limitant aux États occidentaux. Ces sociétés avaient été fondées dans le but de regrouper les évaluateurs en une communauté, afin de stimuler les échanges d'expériences et de donner de la visibilité et de la légitimité à leur expertise auprès des décideurs politiques nationaux ou internationaux. Ce n'est qu'à partir des années 2000 que des sociétés d'évaluation, nationales et régionales, apparurent au Sud[864].

Dans cette partie, nous nous concentrerons sur l'analyse de l'émergence des sociétés nationales d'évaluation au Sud et des sociétés régionales et internationales. Nous montrerons que leur création relève avant tout de facteurs contextuels provenant du système de l'aide au développement, et non de facteurs relevant d'une demande nationale ou d'une volonté des évaluateurs à se rapprocher de leurs pairs d'autres pays.

Le premier facteur contextuel vient des changements dans le système de l'aide qui ont contribué à encourager l'émergence d'organisations de la société civile (OSC) sur les thèmes de l'efficacité du développement. En effet, lorsque la Déclaration de Paris a été adoptée en 2005, des OSC ont critiqué le fait que la ratification n'ait pas assez intégré les acteurs non étatiques du Nord et du Sud. En conséquence, un nouveau forum sur l'efficacité de l'aide a été organisé à Accra en 2008, durant lequel les acteurs de la société civile internationale étaient parties prenantes des négociations. Le lobbying important des OSC internationales lors des négociations de ce forum a permis que la déclaration finale du forum insiste sur la place centrale des OSC dans les processus de développement et sur l'importance de renforcer la collaboration

[864] Une chronologie de la création des sociétés d'évaluation et une carte de leur implantation géographique sont présentées en annexe 4.

entre les organismes gouvernementaux, les OSC et les acteurs du secteur privés[865]. Les donateurs signataires se sont ainsi engagés à : i) améliorer la coordination entre les efforts déployés par les OSC et les programmes gouvernementaux ; ii) renforcer la transparence les OSC sur les résultats de leurs actions ; iii) améliorer l'information sur les activités des OSC ; (iv) et œuvrer avec les OSC à l'instauration de conditions propices à une maximisation de leur contribution au développement[866].

L'importance du rôle de la société civile et du secteur privé dans le développement mondial et la nécessité de créer un partenariat multipartite, dans un contexte où les organismes étatiques de développement ne représentent qu'une minorité des fonds de coopération internationale pour le développement, ont été réaffirmées lors du 4ème Haut Forum sur l'efficacité de l'aide à Busan (Corée du Sud) en 2011. Ce Haut forum s'intitulait d'ailleurs : « Partenariat mondial pour une coopération efficace ». L'accord final stipule que « les OSC jouent un rôle déterminant dans la possibilité pour les personnes de faire valoir leurs droits, dans la promotion des approches fondées sur les droits, dans la définition des politiques de développement et de partenariats et dans leur mise en œuvre. Elles assurent aussi la fourniture de services dans des domaines complémentaires à ceux couverts par l'État »[867]. Les États s'engagèrent à : i) respecter pleinement leurs engagements respectifs en vue de permettre aux OSC de jouer leur rôle d'actrices indépendantes du développement ; ii) créer un environnement cohérent avec les droits internationaux agréés, propice à une maximisation de leur contribution au développement ; iii) à encourager les OSC à suivre des pratiques qui renforcent leur redevabilité et leur contribution à l'efficacité du développement »[868].

Ces deux accords internationaux ont créé un contexte favorable, et même incitateur, à la création de sociétés d'évaluations nationales, régionales et internationales. En effet, ces sociétés contribuent directement à améliorer la bonne gouvernance des politiques de développement en travaillant à diffuser la pratique de l'évaluation à travers le monde. De plus, l'accent mis dans ces déclarations, d'une part sur l'importance de mettre en place des programmes d'aide « ancrés dans les pays » et, d'autre part sur l'efficacité de la coopération « Sud-Sud », encourage les initiatives locales et les partenariats entre acteurs du Sud pour accélérer le développement.

[865] PETIT, Bernard, *La banalisation technocratique de l'aide,* document de travail, Série « Politiques de développement », FERDI, mars 2011.
[866] DAC-OECD, 2008, *op. cit.*
[867] DAC-OECD, *Partenariat de Busan pour une coopération efficace au service du développement,* Busan, 2011.
[868] DAC-OECD, 2011, *op. cit.*

C'est, selon nous, cette légitimation par les organismes publics nationaux et internationaux qui conditionne les capacités et l'influence des réseaux à transformer les politiques publiques.

En plus d'un contexte international qui favorise l'émergence d'OSC sur le thème de l'efficacité du développement et de la bonne gouvernance, le thème de l'évaluation des politiques publiques mobilise de plus en plus les États du Sud et les organismes de développement. Comme nous l'avons étudié dans le chapitre précédent, l'existence de systèmes nationaux d'évaluation est devenue une conditionnalité de l'allocation de l'aide. Ces systèmes sont devenus, dans la pensée des donateurs, un gage de bonne gouvernance et donc de confiance. Cette exigence s'est renforcée à mesure que les accords internationaux sur l'aide au développement mettaient en avant le fait que les États du Sud devaient être davantage meneurs dans les programmes de développement et se les approprier. Pour ce faire, les donateurs les ont incités à mettre en place des systèmes de gestion transparents et redevables des politiques publiques. Ces systèmes permettent aux donateurs de déléguer la gestion des programmes d'aide aux pays bénéficiaires, et donc de les « autonomiser » et de les « responsabiliser », tout en gardant un contrôle étroit sur cette gestion. Les donateurs restent ainsi redevables aux yeux de leurs tutelles financières et sûrs de la préservation de leurs intérêts.

Dans ce contexte, les sociétés nationales d'évaluation se sont multipliées dans les États du Sud, comme une preuve à la fois de l'engagement des États à renforcer leurs capacités en évaluation et de de leur volonté d'inclure la société civile dans ce processus. Alors qu'en 1998, aucun pays en développement ne possédait de sociétés d'évaluation, la quasi-totalité d'entre eux en compte une aujourd'hui[869]. Selon le principe de coopération Sud-Sud, sept sociétés régionales d'évaluation basées dans des pays du Sud ont émergé dans les années 2000, afin de regrouper les sociétés nationales d'évaluation et les évaluateurs du Sud pour qu'ils capitalisent leurs expériences et compétences. Enfin, les donateurs ont concrétisé leur engagement à soutenir les OSC par la création de sociétés transnationales d'évaluation qui soutiennent financièrement et renforcent les capacités des sociétés régionales et nationales. En théorie, un réseau mondial d'évaluateurs aurait donc émergé en une dizaine d'années, avec des relations horizontales à toutes les échelles entre les acteurs du Nord et du Sud[870].

Au-delà d'un contexte international favorable à la montée en puissance des OSC dans le champ de la bonne gouvernance et de l'efficacité de l'aide, une raison à la multiplication des

[869] Excepté, la Tunisie, l'Égypte, l'Algérie, le Tchad, la République Centrafricaine, la Somalie, la Mongolie, l'Irak, les Émirats-Arabes-Unis, Bahreïn, Oman, le Suriname et la Guyane.
[870] Une chronologie de la création des sociétés d'évaluation et une carte de leur implantation géographique sont présentée dans l'annexe 4.

sociétés professionnelles d'évaluation provient du statut particulier de l'évaluation des politiques publiques qui est encore en quête de reconnaissance en tant que profession.

- **L'évaluation, une professionnalisation encore inachevée**

L'évaluation est une pratique récente pour les pays du Sud et pour un large nombre de pays du Nord, notamment les pays d'Europe centrale ou l'Irlande. Dans les pays d'Europe occidentale et d'Amérique du Nord, où la fonction d'évaluation est institutionnalisée depuis plusieurs décennies, il est toujours difficile de qualifier la fonction d'évaluation : est-ce seulement une discipline, une science ou bien une profession ? En effet, même dans les pays qui ont une culture ancienne de l'évaluation des politiques publiques, les débats perdurent, même entre les évaluateurs, pour savoir si l'évaluation peut être considérée comme une profession. Pour Michael Patton, l'évaluation est « une profession exigeante et difficile »[871], pour Ernest House et Yvonna Lincoln, « l'évaluation est une nouvelle profession »[872], alors que pour Peter Rossi, « l'évaluation n'est pas une profession du tout »[873]. D'autres experts, notamment Steve Jacobs et Yves Boisvert[874] présentent les arguments « pour » et « contre » le fait que l'évaluation soit considérée comme une profession, sans trancher la question.

Nous partageons l'hypothèse de Robert Picciottio[875], selon qui l'évaluation est en train de suivre un processus de professionnalisation mais que ce processus n'est pas achevé, ce qui rend la catégorisation de l'évaluation difficile. Cette indétermination du statut de l'évaluation fait que les évaluateurs qui travaillent dans les organismes gouvernementaux nationaux ou internationaux, dans les ONG, dans des cabinets privés ou en tant qu'indépendants sont en recherche de légitimité et de reconnaissance. Nous identifions ici un facteur qui motive les évaluateurs à se regrouper dans des communautés de pairs, afin de gagner en visibilité et d'augmenter leur pouvoir d'influence sur les politiques et l'opinion publiques.

Afin de déterminer dans quelle mesure l'évaluation peut être considérée comme une profession, il convient au préalable de définir ce que nous entendons par profession. À partir

[871] PATTON, Michael, "The challenge of being a Profession", *Evaluation Practice,* vol. 11, n°1, 1990, pp. 45-51.
[872] HOUSE, Ernest, *Professional Evaluation: Social Impact and Political Consequence,* Newbury Park, CA: SAGE, 1993, 205p; LINCOLN, Yvonna, "The ERS standards for program evaluation", *Evaluation and Program Planning,* vol. 8, n°3, 1985, pp. 251-263.
[873] ROSSI, Peter *et al., 2004, op. cit.* pp. 393-394.
[874] JACOBS, Steve, BOISVERT, Yves, "To be or not to be a profession: pros, cons and challenges for evaluation", *Evaluation,* vol. 16, n°4, 2010, pp. 349-369.
[875] PICCIOTTIO, Robert, "The logic of evaluation professionalism", *Evaluation,* vol. 17, n°2, 2011, pp. 165-180.

des travaux des sociologues Geoffroy Millerson[876] et Kate Macdonald[877], nous distinguons plusieurs caractéristiques que doit remplir une profession. Pour qu'une activité devienne une profession, il faut qu'elle soit pratiquée par plusieurs personnes qui partagent un ensemble de codes éthiques, qui détiennent un savoir spécifique acquis par l'éducation ou la formation, qui sont animés d'une conscience de groupe et d'une solidarité envers leurs collègues, qui s'accordent sur des standards de qualité dans la conduite des biens ou des services qu'ils produisent, qui agissent pour le bien public avant de servir leurs intérêts personnels et enfin qui ont un prestige et un statut reconnus par la société. À partir de cette définition, nous allons déterminer sur quels critères l'évaluation répond aux caractéristiques d'une profession et sur lesquels le processus de professionnalisation n'est pas encore achevé.

Concernant l'existence d'un code d'éthique, tous les guides et les politiques d'évaluation produits par les sociétés d'évaluation et les organismes gouvernementaux chargés d'évaluer des politiques publiques en comprennent un. De plus, les manuels d'évaluation comportent généralement un chapitre consacré à l'éthique de l'évaluation[878].

En matière de construction d'un savoir spécifique acquis par l'éducation et la formation, nous observons que la littérature théorique sur l'évaluation est disponible dans un nombre grandissant de pays, notamment grâce à l'action des sociétés d'évaluation qui produisent ou traduisent des manuels d'évaluation dans leurs langues nationales. De plus, le nombre de formations professionnelles et académiques en évaluation augmente, bien qu'il reste limité. En France, par exemple, il en existe une dizaine qui forment environ 200 évaluateurs professionnels par an. C'est bien peu par rapport aux Masters universitaires qui forment à l'audit que nous comptons par dizaines en France et qui forment plusieurs milliers de professionnels par an. Cependant, seuls les États anglo-américains font de l'évaluation une discipline académique à part entière. Nous retrouvons là l'une des limites de la professionnalisation de l'évaluation. En effet, dans la majorité des pays, aucune accréditation n'existe pour garantir les compétences des évaluateurs. Selon Robert Picciottio, « n'importe qui avec une connaissance rudimentaire du jargon évaluatif peut prétendre être un évaluateur. Excepté pour le Canada, les progrès vers l'élaboration d'une certification pour les évaluateurs

[876] MILLERSON, Geoffroy, *The Qualifying Associations: A Study in Professionalization,* Routledge and Kegan Paul Ltd, 1964, 304p.
[877] MACDONALD, Kate M, *The Sociology of the Professions,* Thousand Oaks, CA: SAGE, 1995, 224p.
[878] Tous les guides et politiques des bailleurs de fonds sont accessibles via le site du CAD-OCDE : http://www.oecd.org/derec/guidelines.htm, consulté le 29 novembre 2014.

sont très lents. Les sociétés d'évaluation sont ouvertes à tous les individus ou groupes qui acceptent de payer les frais d'adhésion »[879].

Les sociétés d'évaluation contribuent à développer ou à renforcer un sentiment d'appartenance à un même groupe chez les évaluateurs en créant une plateforme d'échange d'expériences, d'apprentissage mutuel et de partage de bonnes pratiques. Toutefois, les organigrammes des sociétés d'évaluation montrent qu'elles sont fragmentées en groupes sectoriels (aide au développement, politiques de santé, politiques d'éducation, etc.) ou thématiques (évaluation d'impact, évaluation participative, etc.), qui communiquent relativement peu entre eux, excepté lors des conférences plénières nationales, régionales ou internationales organisées une à deux fois par an. Par exemple, la Société française d'évaluation est constituée de 13 groupes sectoriels et thématiques. Cette structuration suit le schéma selon lequel la culture de l'évaluation des politiques publiques s'est diffusée dans les États, c'est-à-dire par secteurs de politiques publiques. Nos échanges avec des évaluateurs, en France et à l'étranger, montrent que les évaluateurs internes à des ministères se sentent plus appartenir au secteur professionnel de leur ministère (santé, éducation, finances, etc.) qu'au groupe professionnel des évaluateurs. De même pour les consultants indépendants qui pratiquent souvent l'évaluation comme une activité supplémentaire à leur profession principale, le plus souvent dans la recherche ou l'enseignement universitaire. Les évaluateurs travaillant dans des cabinets d'évaluation se sentent, quant à eux, appartenir au groupe professionnel des évaluateurs. Toutefois, ces derniers ont tendance à beaucoup différencier l'évaluation de l'aide au développement de celle des politiques publiques nationales, estimant que ces deux activités demandent des compétences très distinctes. Ils soulignent leurs difficultés à travailler avec les organismes d'aide, souvent en raison des compétences demandées ou de trop faibles rémunérations[880].

Concernant l'existence de standards de qualité, l'évaluation est une activité très normée qui dispose d'un ensemble de standards reconnus internationalement, formulés par le CAD-OCDE en 1991. Ils s'appliquent à l'aide au développement mais ont été reproduits par les sociétés d'évaluation, ainsi que par les organismes étatiques qui les ont utilisés dans le cadre leurs chartes ou politiques d'évaluation pour d'autres secteurs de politiques publiques ou générales. Cela tend à remettre en cause la différence soulevée par certains évaluateurs entre l'évaluation de l'aide au développement et l'évaluation des autres politiques publiques

[879] PICCIOTTIO, 2011, *art. cit.* p. 172.
[880] Entretiens informels menés dans le cadre de stages et de missions de consultance au sein de cabinets français d'évaluation des politiques publiques en 2011 et 2013.

nationales. Les sociétés d'évaluation sont également nombreuses à avoir repris les standards de qualité de la société américaine d'évaluation[881], desquels le CAD-OCDE s'est lui-même inspiré dans l'élaboration de ses standards. Toutefois, tous les organismes publics sont loin d'avoir formalisés des standards et politiques d'évaluation et les unités d'évaluation de ces organismes restent souvent peu spécialisées. De plus, dans les pays du Sud, les évaluateurs sont souvent des experts indépendants peu formés à l'évaluation et qui ne sont pas protégés par une structure de la pression de leurs commanditaires[882]. Les sociétés d'évaluation jouent un rôle important sur cet aspect, dans la mesure où elles contribuent à sensibiliser les producteurs et les utilisateurs des évaluations aux standards de qualité auxquels doivent répondre les évaluations pour être considérées comme étant indépendantes, rigoureuses et donc professionnelles. Elles permettent également de renforcer les capacités en évaluation de leurs membres, afin qu'ils améliorent la qualité de leurs travaux. Enfin, elles constituent une structure de rattachement pour les évaluateurs indépendants.

La littérature théorique sur la professionnalisation de l'évaluation est consensuelle sur le fait que, dans les pays occidentaux, l'émergence d'une offre en évaluation répond clairement à une demande des institutions publiques et de la société civile pour une gestion plus transparente, redevable et efficace des politiques publiques. L'évaluation aurait ainsi comme principe fondateur le bien public. Un évaluateur des Nations unies estime par exemple que « *l'intérêt public est une préoccupation essentielle parmi les évaluateurs. Les évaluateurs se caractérisent souvent par leur engagement dans leur carrière malgré de faibles revenus par rapport aux activités proches (audits, contrôle, etc.), leur préoccupation de fournir un travail de qualité et une volonté constante d'améliorer leurs compétences* »[883].

Enfin, nous allons nous intéresser au prestige de l'évaluation dans la société, un des points les plus faibles de la professionnalisation de l'évaluation. La distinction entre l'évaluation, l'audit, le contrôle et la recherche en sciences sociales apparaît souvent floue aux personnes que nous avons interrogées à ce sujet au sein d'agences d'aide[884]. Selon Peter Rossi, l'évaluation est « un champ amorphe, diversifié et hétérogène avec des profils d'évaluateurs très variés allant de l'économiste à l'ingénieur, en passant par le chercheur en sciences

[881] DAC-OECD, 2013, *op. cit.*
[882] PICCIOTTIO, 2011, *art. cit.*
[883] Entretien avec un évaluateur des Nations unies, 6 octobre 2014.
[884] Constat issu de discussions informelles avec des opérationnels de l'aide au développement lors de missions de consultance au sein de l'AFD, du MAE, de la Cour des comptes.

sociales »[885]. Cette situation conduit au manque de reconnaissance de la discipline et à ce que « peu de personnes dans l'évaluation ont réussi à atteindre des postes à responsabilité et des récompenses en grimpant les échelons hiérarchiques à partir de leur expérience dans les unités d'évaluation »[886]. Le statut et le prestige des évaluateurs sont affaiblis par la concurrence des auditeurs, des chercheurs en sciences sociales et des contrôleurs des finances publiques. D'autant plus que le nombre de professionnels de l'évaluation est réduit par rapport à ces groupes professionnels. Aux États-Unis par exemple, les entreprises d'audit comptent environ 100 000 professionnels, soit au moins dix fois plus que le nombre d'évaluateurs[887]. Ce différentiel fragilise le statut des évaluateurs qui devront parvenir à augmenter leur masse pour pouvoir rentrer en compétition avec les autres groupes professionnels[888]. Un autre facteur qui nuit au statut des évaluateurs est le fait que la plupart d'entre eux travaillent en même temps en tant qu'enseignants, chercheurs ou consultants en sciences sociales et ne dédient qu'une partie de leur temps à l'évaluation, ne se qualifiant alors pas eux-mêmes d'évaluateurs.

Nous concluons donc de cette étude que l'évaluation est en cours de professionnalisation. Nous pensons que c'est la volonté des évaluateurs de professionnaliser leur fonction qui les a poussés à mettre en place des sociétés professionnelles d'évaluation. Nous allons désormais vérifier cette hypothèse en analysant les raisons pour lesquelles les sociétés d'évaluation ont été créées et quel a été leur processus de création.

- **Des motivations politiques inspirées par les organisations multilatérales**

Cette analyse de l'émergence de sociétés d'évaluateurs au Sud se fonde sur des entretiens menés en 2013 auprès de responsables de ces sociétés et d'un questionnaire réalisé par l'IOCE dont nous avons analysé et compilé les réponses pour obtenir une image globale de ce phénomène[889].

Dans la littérature en relations internationales, l'émergence des réseaux est souvent en partie expliquée par l'organisation de conférences thématiques à l'échelle régionale ou

[885] ROSSI *et al.*, 2004, *op. cit.* p. 395. La diversité de ces profils est illustrée dans l'annexe 10, dans laquelle nous avons retracée le parcours académiques et professionnels d'une quinzaine d'évaluateurs que nous avons interviewés dans le cadre de notre thèse.
[886] *Ibid.*
[887] Données recueillies à partir de la base de données du réseau transnational d'évaluation IOCE à laquelle nous avons pu accéder. Nous avons consolidé les données de cette base pour obtenir des informations sur l'état actuel des sociétés d'évaluation à travers le monde.
[888] PICCIOTTIO, 2011, *art. cit.*
[889] Les questionnaires individuels par organisation sont consultables sur le site de l'IOCE : http://www.ioce.net/en/members.php, consulté le 1er décembre 2014.

nationale. Margaret Keck et Kathryn Sikkink citent, par exemple, « la tenue de conférences ou autres évènements internationaux qui froment des arènes pour créer et renforcer les réseaux »[890]. Peter Haas, quant à lui, identifie l'intensification des échanges d'idées, notamment *via* des conférences et des collaborations internationales, comme cause de l'émergence de communautés épistémiques[891]. Nous avons pu observer ce schéma dans la création de certaines sociétés nationales d'évaluation du Sud (au Nigéria, au Mali, au Vietnam), où la décision de créer une société d'évaluation est venue de professionnels de l'évaluation, après qu'ils ont participé à une ou plusieurs conférences sur l'évaluation organisées par des sociétés régionales ou internationales d'évaluation. Leur initiative fut ensuite soutenue financièrement et logistiquement par les Nations unies (souvent l'UNICEF). Pour d'autres sociétés d'évaluation (au Cambodge, en Équateur, au Népal, en Ukraine), le processus est le même, trouvant son origine dans la participation de professionnels de l'évaluation à des conférences sur l'évaluation, à la différence que ces derniers se sont ensuite organisés en associations sans demander l'appui d'organisations internationales.

Un autre schéma identifié par la théorie des relations internationales est l'impulsion plus directe par des organisations internationales. Peter Haas observe que « les idées d'une communauté épistémique peuvent prendre racine dans une organisation internationale, après quoi elles se diffusent aux États par l'action des acteurs qui ont été influencés par ces idées »[892]. Cette dynamique a été très courante dans le cadre des sociétés d'évaluation, où c'est l'action concertée d'organismes des Nations unies - dans 90 % des cas de l'UNICEF - et des gouvernements qui a mené à leur création (au Burkina Faso, en Côte d'Ivoire, en Mauritanie, aux Niger, aux Philippines). Dans d'autres cas, le gouvernement n'était pas demandeur et l'initiative provient de professionnels de l'évaluation (en Éthiopie, au Sri Lanka, au Honduras) ou de chercheurs (en Égypte, en Thaïlande, Ouganda) qui ont trouvé une inspiration ou un soutien dans les organisations internationales, sans pour autant avoir participé à des séminaires de sociétés internationales ou régionales d'évaluation. Encore une fois, le nom de l'UNICEF revient dans la quasi-totalité des récits. Enfin, dans certains cas, l'initiative était purement nationale et semblait peu influencée *a priori* ou *a posteriori* par les organismes internationaux de développement (au Chili, au Ghana, en Indonésie, au Pérou, au Maroc, etc.).

---

[890] KECK, Margaret, SIKKINK, Kathryn, "Transnational advocacy networks in international and regional politics", *International Social Science Journal*, vol. 51, issue 159, mars 1999, pp. 89-101.
[891] HAAS, Peter, "Epistemic Communities and International Policy Coordination", *International Organization*, vol. 46, n°1, hiver 1992, pp. 1-35.
[892] HAAS, 1997, *op. cit*, p. 17.

Cette typologie fait ressortir le rôle important de l'UNICEF dans la création de sociétés nationales d'évaluation dans les pays du Sud. Nous nous sommes alors questionné quant à la raison de l'implication de cet organisme sur cette thématique.

Selon plusieurs de nos interviewés[893], c'est Mahesh Patel, *social policy advisor* à l'UNICEF, qui a été à l'origine de l'engagement de l'organisation dans l'appui à la création de sociétés nationales et régionales d'évaluation. Ils décrivent Mahesh Patel comme quelqu'un qui croyait profondément à la capacité des sociétés professionnelles d'évaluation d'améliorer la culture de la transparence et de l'évaluation dans les pays du Sud. Durant toute sa carrière, il a dédié beaucoup d'énergie à encourager des professionnels de l'évaluation et des fonctionnaires locaux à mettre en place des sociétés d'évaluation en Afrique, sa zone géographique d'affectation. Il a notamment été l'un des fondateurs de l'Association africaine d'évaluation (AFREA). Son engouement a influencé ses collaborateurs, spécifiquement Marco Segone qui suivit son exemple et encouragea les évaluateurs à créer des sociétés professionnelles dans tous les pays où il était affecté, de l'Europe centrale à l'Asie du Sud, en passant par l'Afrique. Il y a sept ans, Marco Segone a obtenu un poste à New York au siège de l'UNICEF, au sein du département de l'évaluation, et a continué son travail. Il a sensibilisé le personnel du siège à la question du renforcement des capacités en évaluation et du rôle des sociétés professionnelles d'évaluation à cet effet. Il a réussi à convaincre la direction de l'importance d'impliquer l'UNICEF sur cette thématique. Depuis, l'UNICEF collabore étroitement avec les sociétés d'évaluation, notamment, au niveau international *via* le réseau international d'évaluateurs IOCE dont Marco Segone a encouragé la création.

À travers cette histoire, nous remarquons que l'émergence à l'UNICEF de la problématique du soutien aux sociétés professionnelles d'évaluation a reposé sur quelques individus seulement. Nous retrouvons ce poids des personnalités dans la grande majorité des récits que nous avons lus et entendus sur la naissance des sociétés d'évaluation. À chaque fois, deux ou trois individus sont cités comme ayant porté le projet jusqu'à ce qu'il aboutisse, sur une période plus ou moins longue selon les contextes. Nous en déduisons que les réseaux, qu'ils soient dédiés à l'évaluation ou à d'autres enjeux, sont avant tout une histoire d'individus qui croient en une cause et qui vont se mobiliser pour la politiser.

Ce soutien de l'UNICEF peut apparaître bénéfique aux sociétés d'évaluation car il leur a permis de se former rapidement, en mettant à leur disposition des moyens financiers, logistiques

[893] Entretien avec le directeur d'un cabinet d'évaluation, 20 novembre 2013 ; entretien avec la co-responsable du réseau d'évaluateurs Evalpartners, entretien Skype, 22 janvier 2014.

et d'expertise importants. Il a en grande partie permis la multiplication des sociétés d'évaluation au Sud durant les années 2000. Toutefois, nous estimons que ce soutien handicape ces sociétés sur le long terme, dans la mesure où une fois que les organisations internationales se désengagent, elles n'ont souvent pas les moyens de continuer à mobiliser autant de ressources humaines et financières et de mettre en place des évènements d'envergure nationale ou régionale. Natalia Koshelava témoigne à ce sujet : « *certaines sociétés professionnelles d'évaluation ont été créées par les donateurs. Un schéma type de développement d'une société d'évaluation est le suivant : pendant quelques années, l'UNICEF va financer des conférences nationales sur l'évaluation dans le pays, ainsi que la venue des participants intéressés. Les conférences sont organisées dans des hôtels cinq étoiles et des intervenants internationaux renommés sont conviés à participer. Les donateurs vont alors inciter les participants à créer une société nationale d'évaluation et les assister dans cette tâche. Une fois que la société est fonctionnelle, les donateurs se retirent et disent à ses dirigeants qu'ils vont devoir être autonomes et gérer l'association eux-mêmes. Le problème est que les donateurs ont créé des attentes fortes chez les membres de l'association en les faisant participer à des évènements de grande ampleur et de très bonne qualité. Or, les sociétés nationales d'évaluation au Sud n'ont pas les moyens de reproduire le même type d'évènements, car elles sont gérées par des bénévoles et qu'il est impossible de générer des ressources financières par les frais d'inscription, qui, s'ils sont trop élevés, deviennent inabordables pour les populations du Sud* »[894].

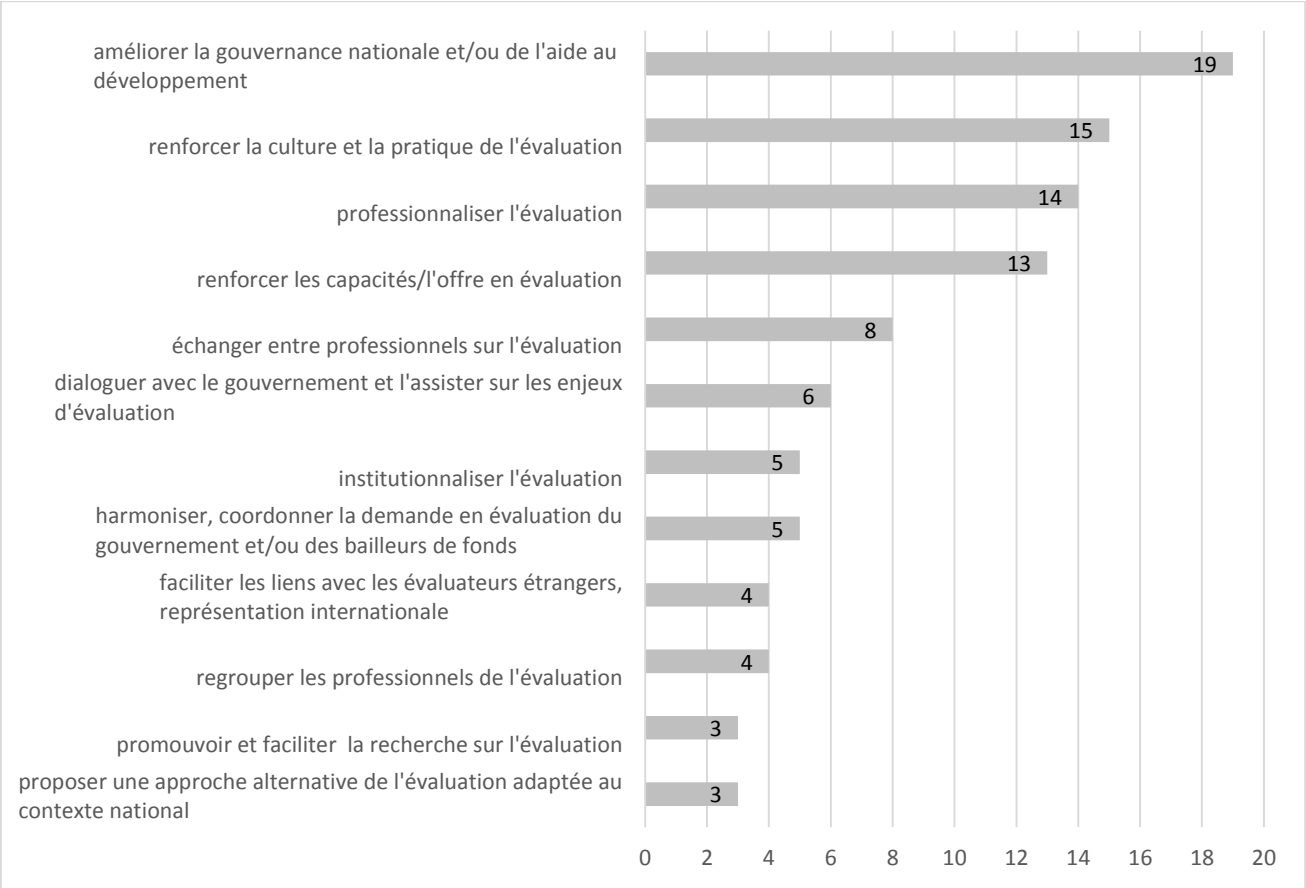
Un soutien trop important d'organisations multilatérales peut ainsi affaiblir les réseaux nationaux d'évaluateurs qui survivent mal à leur désengagement car leurs membres sont découragés par le manque de moyens et de soutien politique qui s'en suit. Nous remarquons que les sociétés d'évaluation les plus actives sont souvent issues d'une initiative locale, avec des évaluateurs qui se regroupent car ils ont la volonté personnelle de créer une association (Ukraine, Nigéria, Pérou) quelles que soient les difficultés rencontrées. En effet, ces sociétés professionnelles ont eu l'habitude de fonctionner, dès le départ, avec des moyens limités reposant sur l'implication forte d'un groupe de bénévoles.

Au-delà des alliances d'acteurs qui ont permis la création des sociétés d'évaluation au Sud, nous avons également analysé les motivations qui les ont poussés à entreprendre de telles

[894] Entretien avec une consultante en évaluation à Process Consulting Company en Russie, co-responsable du réseau d'évaluateurs Evalpartners, 22 janvier 2014.

initiatives. À partir d'une enquête réalisée par l'IOCE[895], nous avons établi une base de données recensant les différentes raisons évoquées par les représentants des sociétés d'évaluation du Sud interrogés. Notre analyse a donné les résultats suivants:

Graphique 1 : Motivations à la création de sociétés d'évaluation dans les États du Sud



Source : questionnaire de l'IOCE aux sociétés d'évaluation, 2013

Ce graphique fait ressortir quatre types de raisons à la création de sociétés nationales d'évaluation dans les pays du Sud. Les premières sont l'amélioration de la gouvernance nationale (hausse de la transparence, ouverture envers la société civile, meilleure efficacité des politiques publiques, etc.) et de la gestion des programmes d'aide au développement (meilleure participation des acteurs locaux, alignement[896] des donateurs, etc.). Ces motivations relèvent d'une volonté de changement de la culture politique nationale. Il nous a semblé étonnant de les voir apparaître en premier lieu, étant donné qu'elles semblent très ambitieuses pour des sociétés d'évaluation qui ne regroupent souvent que quelques dizaines de membres. Cependant, la

[895] En 2013, l'IOCE a réalisé une enquête auprès des sociétés d'évaluation pour avoir une vision de leur fonctionnement. Une grande partie des résultats de cette enquête n'a pas été exploitée. Nous les avons utilisés et consolidés, pour en tirer des informations sur la genèse, le fonctionnement et les résultats des sociétés d'évaluation.
[896] Le terme "alignement" est défini dans le lexique en annexe 7.

littérature théorique montre que ce sont des motivations fréquentes à la création de réseaux, ces derniers émergeant souvent avec l'objectif de pallier un manque des gouvernements[897]. Ici, il s'agirait d'un manque de transparence et de redevabilité dans la gestion des politiques publiques et des programmes d'aide. La création d'un réseau permet à un groupe d'acteurs de faire valoir ses compétences et ses revendications avec plus de force et de visibilité, ainsi que de faire pression sur le gouvernement pour qu'il modifie ses pratique

Le deuxième type de raison tend à ce que les membres des sociétés d'évaluation veulent que les ministères et les donateurs coordonnent leur demande d'évaluation et que les premiers institutionnalisent l'évaluation dans des textes législatifs. Ils espèrent également que leur regroupement en réseau leur permettra d'accéder aux travaux du gouvernement sur la planification et la mise en œuvre des évaluations, de rendre leur expertise plus visible aux yeux du gouvernement, ce qui pourra créer une demande d'évaluation de la part de ce dernier. Les évaluateurs sont donc à la recherche d'une influence politique dans leur pays. Les chercheurs ayant travaillé sur la sociologie des réseaux montrent que les réseaux apparaissent dans des contextes d'incertitude, où leurs membres vont jouer un rôle d'informateurs auprès des décideurs politiques[898].

La troisième motivation récurrente de la formation de réseaux d'évaluateurs au Sud est le renforcement des capacités en évaluation, qui passe par la mise en place de formations, d'ateliers de travail, de séminaires, etc. et la professionnalisation de l'évaluation, qui s'effectue par la création de diplômes universitaires et d'accréditations, la formalisation de normes de qualité et d'éthique et la création d'un sentiment d'appartenance à une communauté professionnelle.

Enfin, la quatrième raison, qui ressort le plus rarement, est la mise en réseau de professionnels au niveau national et international. Elle semble pourtant refléter la réalisation la plus immédiate des associations de professionnels.

Les motivations les plus évoquées sont donc les plus politiques et de long terme, qui touchent au développement d'une demande d'évaluation, alors que les fonctions que pourraient plus facilement remplir les sociétés d'évaluation, qui touchent à la création d'une offre d'évaluation, sont moins citées. Cela peut s'expliquer par le fait que ces sociétés ont été créées pour pallier les manquements des programmes de renforcement des capacités en évaluation des donateurs qui portaient surtout sur l'offre d'évaluation et très peu sur la demande. Ainsi, les

[897] HAAS, 1997, *op. cit* ; KECK, SIKKINK, 1998, *op. cit.*
[898] COLONOMOS, Ariel, 2001, *op. cit.*

sociétés d'évaluation du Sud ont été incitées par les donateurs à se consacrer au renforcement d'une culture et d'une demande de l'évaluation au niveau du gouvernement et de la société civile et donc à l'amélioration de la gouvernance de leur pays, plus qu'à agir sur l'offre d'évaluation.

Les sociétés d'évaluation au Sud sont étroitement liées à des enjeux de développement. Dès leur création, elles sont vouées à être des acteurs du système international de l'aide. Leur mise en réseau international a été permise par la création de sociétés internationales ou régionales d'évaluation les chaperonnant. Aujourd'hui, chaque continent ou sous-continent du monde a sa société régionale d'évaluation et les organisations multilatérales de développement ont créé des sociétés d'évaluation à vocation internationale (IDEAS, IOCE, Evalpartners). Enfin, des sociétés internationales thématiques ont émergé, par exemple sur la prise en compte des enjeux environnementaux ou du genre dans les évaluations, ou encore sur les évaluations expérimentales[899].

Comme pour les sociétés nationales, nous avons étudié, à partir du questionnaire de l'IOCE, les raisons de la création des sociétés régionales et internationales d'évaluation. Tout d'abord, elles ont vocation à soutenir les sociétés nationales d'évaluation. En effet, en regroupant un grand nombre d'évaluateurs de nationalités multiples au sein d'un même espace, elles compensent le manque de capital humain dont souffrent la majorité des sociétés d'évaluation du Sud et génèrent des échanges d'expériences et d'expertise sur l'évaluation. Les sociétés régionales et internationales aspirent ainsi à faire émerger une communauté transnationale d'évaluateurs qui partagent une éthique commune de l'évaluation dépassant les différences contextuelles et culturelles.

Ce soutien permet également de dépasser l'isolement des sociétés nationales d'évaluation vis-à-vis des instances gouvernementales de leur État d'origine. Le soutien de sociétés régionales et internationales augmente leur légitimité et leur donne plus de pouvoir au niveau national. Margaret Keck et Kathryn Sikkink analysent ce processus qu'elles qualifient « d'effet boomerang » : « beaucoup de réseaux internationaux lient des activistes de pays développés à d'autres de pays moins développés. Ce type de mise en réseau a souvent pour objectif d'influencer le comportement des États. Lorsque les liens entre l'État et les acteurs des réseaux nationaux sont faibles, ces derniers peuvent directement chercher des alliés à l'international pour essayer de faire pression sur leur État de l'extérieur »[900]. Si l'on suit cette analyse, les

[899] Voir carte des sociétés professionnelles d'évaluation en annexe 4.
[900] KECK, Margaret, SIKKINK, Kate, 1998, *op. cit.* p. 71.

organisations multilatérales auraient créé des sociétés d'évaluation internationales et régionales, afin d'aider les réseaux nationaux à influencer la politique de leur gouvernement vers une meilleure gouvernance. Nos recherches nous ont permis de confirmer cette hypothèse. Lorsque les responsables des sociétés internationales d'évaluation sont interrogés sur les motivations de leur action, ils répondent avoir voulu « soutenir la gestion axée sur les résultats dans la gouvernance des politiques publiques des États du Sud », « promouvoir la gouvernance démocratique dans le monde », ou encore « renforcer les capacités des réseaux d'évaluateurs à influencer les politiciens, l'opinion publique et les autres parties prenantes des politiques publiques dans leurs pays ».

Les représentants des sociétés régionales et internationales d'évaluation citent également l'objectif de rendre plus équitables les relations entre les évaluateurs du Nord et du Sud. Ils aspirent à professionnaliser et valoriser l'expertise du Sud pour qu'elle ait un statut équivalent à celle du Nord sur le marché international de l'évaluation, notamment dans le cadre de l'aide au développement. Par extension, les sociétés internationales d'évaluation visent à participer à l'entreprise globale d' « égalitarisation » des relations entre donateurs et bénéficiaires de l'aide.

En termes de logique d'acteurs, nous retrouvons le même schéma identifié pour les sociétés nationales, c'est-à-dire une impulsion donné par un petit groupe de professionnels de l'évaluation, soutenu par une organisation multilatérale. Par exemple, une ancienne Présidente de l'Association africaine d'évaluation (AFREA) raconte que ce réseau « *est né en 1999 à la suite d'une réunion entre les responsables des six sociétés d'évaluation actives en Afrique à cette époque. Cette réunion avait été organisée par l'UNICEF, qui a ensuite apporté les moyens financiers et en expertise nécessaires à la concrétisation de l'initiative. Des acteurs de la société civile internationale ont également soutenu l'AFREA, notamment les ONG CARE et Catholic Relief Service. Un des effets les plus visibles de la création de l'AFREA a été la multiplication des sociétés africaines d'évaluation, qui ont émergé dès les années suivantes. Cela est dû au plaidoyer effectué par l'AFREA envers les évaluateurs nationaux pour qu'ils mettent en place des sociétés nationales d'évaluation* »[901]. Aujourd'hui, l'AFREA regroupe 20 sociétés d'évaluation qu'elle soutient et guide par la création de normes sur l'évaluation, par des actions de renforcement des capacités et en encourageant les évaluateurs nationaux à plaider pour une meilleure gouvernance dans leurs pays, afin de rendre leurs environnements plus favorables au développement de l'évaluation des politiques publiques.

[901] Entretien avec une ex présidente de l'AFREA, 22 mai 2012

En ce qui concerne la société internationale d'évaluation IOCE, sa création est le résultat du volontarisme de quelques personnes au sein de l'UNICEF, notamment Marco Segone, qui croyaient en la capacité des sociétés d'évaluation de professionnaliser et de systématiser la pratique de l'évaluation à travers le monde. L'IOCE reçut le soutien de la Banque mondiale, de *l'International Fund for Agriculture and Development* et de plusieurs fondations privées (dont une implication importante de la *Bill Gates Fundation*). Aujourd'hui, l'IOCE est elle-même à l'origine de la création d'un nouveau réseau international d'évaluateurs, Evalpartners, qui a pour but d'améliorer le partage d'expérience et l'équité entre les évaluateurs du Nord et du Sud[902].

Nous voyons donc que la création de réseaux d'évaluateurs transnationaux entraîne l'émergence de nouveaux réseaux. Nous sommes dans une période d'émulation mondiale où les évaluateurs cherchent à se regrouper en réseaux plus ou moins larges pour exister en tant que communauté professionnelle. Ce constat rejoint l'analyse de Margaret Keck et Kathryn Sikkink : « les réseaux nourrissent les réseaux ; quand le réseautage devient un répertoire d'action qui se diffuse internationalement, chaque nouvel effort pour créer un réseau international est moins difficile que le précédent »[903].

B. Donner une existence politique aux évaluateurs

Après avoir analysé les processus d'émergence de ces sociétés d'évaluation et leurs ambitions, nous allons désormais étudier quel est leur mode de fonctionnement et quelles sont leurs activités effectives dans l'espace international. Nous étudierons également les rapports de forces internes et externes qui les animent.

- **Des logiques d'acteurs hiérarchisées**

La taille des sociétés nationales d'évaluation est très variable, allant de moins d'une dizaine de membres, souvent dans les pays du Sud (Colombie, Serbie, Tanzanie, Népal, etc.), à plusieurs milliers, le plus fréquemment dans les pays du Nord (États-Unis, Canada), ou bien dans celles ayant une envergure régionale ou internationale. Au total, les sociétés

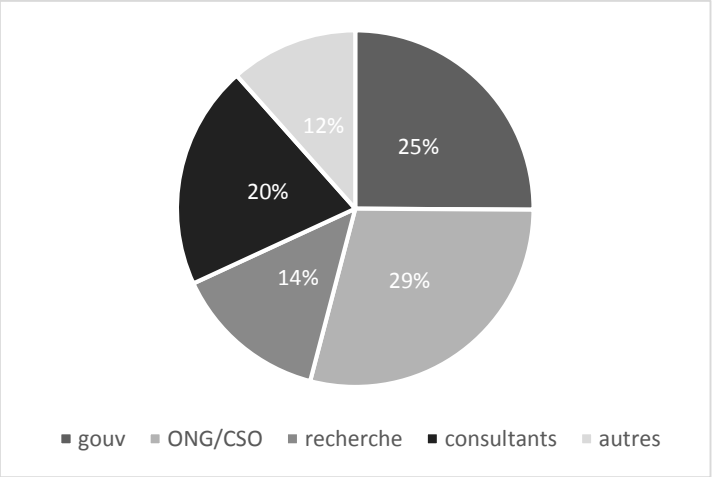
[902] KOSHELEVA, Natalia, SEGONE, Marco, "Evalpartners: Facilitating the Development of a New Model of Voluntary Organization for Professional Evaluation to Support the Development of National Evaluation Capacities", *American Journal of Evaluation*, 2013, pp. 1-6.
[903] KECK, Margaret, SIKKINK, Kathryn, 1999, *art. cit.*

professionnelles d'évaluation regroupent aujourd'hui plus de 34 000 évaluateurs[904], mais les données que nous avons pu rassembler et analyser pour cette étude ne distinguent pas les membres actifs et inactifs des sociétés. De plus, ce total est biaisé à la hausse par le fait qu'un même individu peut adhérer à plusieurs sociétés d'évaluation en même temps. Plusieurs adhésions décomptées peuvent donc, en réalité, représenter un seul et même membre[905]. Toutefois, ce chiffre permet d'apprécier l'augmentation de l'expertise en évaluation à travers le monde et la hausse de l'intérêt porté à cette activité. En effet, les sociétés professionnelles ne comptaient qu'environ 4 000 adhérents jusque dans les années 1990, soit huit fois moins qu'aujourd'hui[906].

Pour comprendre le fonctionnement et la nature des sociétés professionnelles d'évaluation, il nous semble intéressant d'étudier le profil de leurs membres. Nous nous sommes concentrée sur les sociétés d'évaluation des pays du Sud, régionales et internationales, qui rentrent dans le cadre de notre analyse sur l'aide au développement. Tout d'abord, concernant les sociétés nationales d'évaluation au Sud, les données que nous avons pu rassembler donnent la composition suivante :

Graphique 2 : Secteurs d'activités des membres des sociétés nationales d'évaluation au Sud

[904] RUGH, Jim, SEGONE, Marco, Editors, Voluntary Organizations for Professional Evaluation (VOPEs), Learning from Africa, Americas, Asia, Australasia, Europe and Middle East, UNICEF, Evalpartners, IOCE, 2013.
[905] Par exemple, l'association américaine d'évaluation rapport que 15 % de ses adhérents sont des internationaux, c'est-à-dire des citoyens d'autres pays, dont on peut supposer qu'ils sont également membres de leur société nationale d'évaluation en plus d'être membre de l'association américaine d'évaluation, et peut être aussi d'autres sociétés régionales et internationales.
[906] RUGH, Jim, SEGONE, Marco, 2013, *op. cit.*



Source : enquête de l'IOCE sur les sociétés d'évaluation, 2013

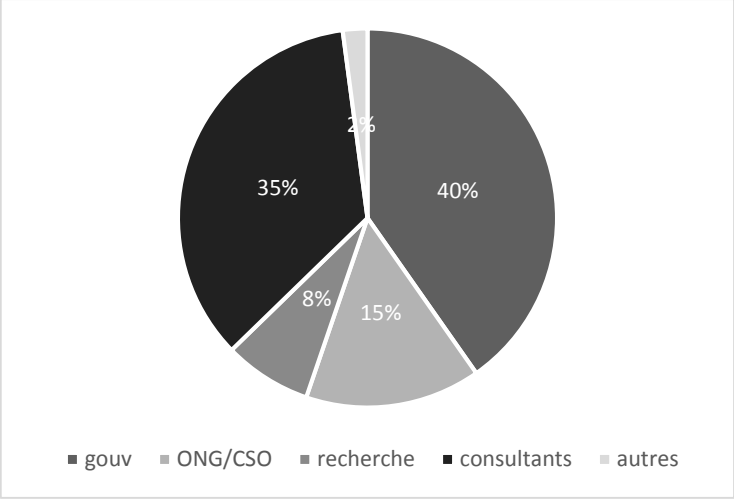
Les sociétés nationales d'évaluation au Sud sont composées à 29 % de représentants de la société civile. Il s'agit d'individus appartenant à des organisations non gouvernementales opérant sur les thématiques de la bonne gouvernance, de la lutte contre la corruption, de la hausse de la transparence, ou encore sur des thématiques socio-économiques : lutte contre la pauvreté, préservation de l'environnement, égalité hommes-femmes, etc. Ces acteurs ne sont souvent pas des évaluateurs à proprement parler, mais des personnes qui s'intéressent à l'évaluation comme un outil d'amélioration de l'efficacité des politiques publiques. Les représentants des gouvernements forment la deuxième catégorie de membres des sociétés d'évaluation au Sud (25 %). Certains d'entre eux font partie d'unités d'évaluation des ministères. Toutefois, étant donné que la fonction d'évaluation reste peu développée dans la majorité des États du Sud, la plupart ne travaillent pas directement dans l'évaluation, mais ont été confrontés *via* des programmes d'aide internationaux et qui ont ensuite développé un intérêt pour cette thématique. Les évaluateurs du secteur privé n'arrivent qu'en troisième position (20 %). Cela s'explique par la faible expertise présente dans les pays du Sud. Les cabinets de consultance en évaluation des politiques publiques sont rares et les consultants indépendants sont éparpillés sur le territoire. Ils travaillent davantage auprès des organisations internationales qu'auprès de leurs gouvernements. Le milieu de la recherche représente 14 % des membres et est majoritairement composé de chercheurs ou des doctorants en sciences sociales. La catégorie « autre » peut comprendre les représentants d'organisations internationales, les médias, les acteurs du secteur privé, etc.

Nous observons donc que les professionnels de l'évaluation représentent une minorité des membres des sociétés nationales d'évaluation au Sud. Ces sociétés sont très ouvertes vers l'extérieur et acceptent un panel diversifié de membres. Cela leur permet de se développer, de

gagner des ressources et de diffuser plus efficacement les valeurs et les pratiques qu'elles promeuvent. Elles visent, avant tout, à construire une demande d'évaluation et cherchent donc des appuis dans le milieu de la recherche, au sein de la société civile et surtout des instances politiques pour avoir plus de relais d'influence pour faire émerger cette demande.

Lorsque nous regardons la composition des sociétés régionales d'évaluation implantées au Sud, les résultats sont différents :

Graphique 3 : Secteurs d'activités des membres des sociétés régionales d'évaluation du Sud



Source : enquête de l'IOCE sur les sociétés d'évaluation, 2013

Les représentants des gouvernements constituent la majorité des membres de ces sociétés (40 %). La part des consultants en évaluation, quant à elle, double (35 %). Cet écart s'explique en partie par l'adhésion des évaluateurs provenant d'États où il n'y a pas de sociétés nationales pour les représenter, bien que cela soit de plus en plus rare. Une seconde explication, plus significative, est qu'un bon nombre d'évaluateurs indépendants des pays du Sud deviennent directement membres des sociétés régionales d'évaluation, sans passer par l'adhésion aux sociétés nationales. En effet, certains de ces évaluateurs, faute d'une demande nationale en évaluation suffisante, travaillent très largement à l'international, auprès d'organismes de développement gouvernementaux et non gouvernementaux. De plus, nous pouvons y voir le reflet de « l'effet boomerang » développé par Margaret Keck et Kathryn Sikkink[907], selon qui les « activistes » peuvent trouver plus efficace d'agir de l'extérieur, dans des réseaux régionaux ou internationaux, pour entraîner un changement au sein de leur gouvernement, lorsque les

[907] KECK, Margaret, SIKKINK, Kathryn, 1998, *op. cit.*

réseaux nationaux ne sont pas opérationnels ou qu'ils sont coupés des instances politiques. Nous constatons, en effet, que les réseaux régionaux et transnationaux d'évaluation se définissent en tant que fournisseurs de ressources et d'influence aux acteurs locaux : « à travers les partenariats et les réseaux, les sociétés d'évaluation individuelles bénéficient d'opportunités d'échanges de bonnes pratiques et de connaissances, ainsi que de coopération. Ces coopérations leur permettent d'étendre leur sphère d'influence *via* des partenariats Sud-Sud ou Nord-Sud. Les objectifs qui sont difficiles à atteindre par les sociétés nationales d'évaluation lorsqu'elles sont isolées deviennent alors beaucoup plus simples à réaliser dans le cadre de partenariats soutenus internationalement, non seulement par les autres sociétés d'évaluation, mais également par les organisations de développement, telles que les Nations unies »[908].

Enfin, nous remarquons une baisse considérable des membres issus de la société civile par rapport aux sociétés nationales d'évaluation. En effet, les OSC ont un ancrage et une influence nationaux et que leur action n'entre pas dans un cadre international. Les représentants des OSC adhèrent souvent aux sociétés nationales d'évaluation car ils en partagent les intérêts et les valeurs, mais ne sont pas des membres très actifs. De plus, ils n'ont pas forcément les moyens de participer aux évènements organisés par les réseaux régionaux. Les représentants du gouvernement et surtout les évaluateurs indépendants participent plus activement et ont plus d'intérêt à aller chercher un soutien international ou régional pour renforcer leur action au niveau national. De plus, apparaître dans les évènements et les listes de diffusion des réseaux d'évaluateurs régionaux et internationaux parrainées par des organisations multilatérales est une vitrine professionnelle et/ou politique pour les ces acteurs. Les évaluateurs se font connaître par leurs pairs et les organismes de développement et les personnalités politiques affichent sur la scène internationale leur engagement à développer l'évaluation et donc à améliorer leur gouvernance.

Après avoir étudié la composition des sociétés d'évaluation, nous allons analyser comment elles interagissent et s'organisent. Si l'émergence des réseaux transnationaux a été théorisée par les chercheurs en relations internationales, peu de travaux modélisent leur fonctionnement et les instruments dont ils disposent pour atteindre leurs objectifs.

D'après Ariel Colonomos, les réseaux internationaux sont le symbole de « la revanche de l'informel sur le formel (…) L'informalité est souvent un gage de contournement de l'État

[908] RUGH, Jim, SEGONE, Marco, 2013, *op. cit.*

et fait accéder les individus à des formes inédites d'action »[909]. L'idée que cette informalité serait revendiquée comme un atout par les réseaux est en partie remise en cause par l'étude que nous avons menée. L'enquête réalisée par l'IOCE en 2013 sur les sociétés d'évaluation montre que 36 % de ses sociétés sont informelles, que 15 % ont créé une charte et un statut, mais ne sont pas officiellement reconnues par leurs gouvernements, tandis que 49 % sont légalement reconnues[910]. Selon les membres des réseaux d'évaluation dont nous avons pu recueillir les témoignages[911], le passage d'un réseau informel à une association formelle est marquée par l'adoption de lois, d'une charte ou d'une constitution, la reconnaissance par le ou les gouvernements du ou des États représentés et l'institutionnalisation du réseau avec la création d'un conseil de direction. Des recherches plus approfondies à partir de cette enquête nous ont permis de constater que la grande majorité des sociétés d'évaluation informelles se situent dans les pays du Sud (seules deux sociétés d'évaluation au Nord sont informelles). Au premier abord, cette observation irait dans le sens de la thèse d'Ariel Colonomos, car les sociétés nationales d'évaluation ont tendance à être informelles dans les pays et les régions les plus touchés par des problèmes de manque de transparence et de corruption, soit ceux où l'environnement administratif est le moins favorable au développement de l'évaluation. L'informalité pourrait donc bien être une stratégie des membres des réseaux d'évaluateurs pour contourner l'État et exercer une pression externe et diffuse sur ce dernier en faveur d'une amélioration de la culture de l'évaluation.

Toutefois, notre recherche a démontré que, dans la plupart des cas, l'informalité est forcée et est vue comme un handicap par les membres des réseaux d'évaluateurs. L'ex-présidente de l'association d'Europe centrale d'évaluation *l'International Programme Evaluation Network* (IPEN), déclare, par exemple, que « *si l'IPEN reste un réseau régional informel, c'est avant tout à cause des législations des pays membres qui empêchent la formation d'associations formelles* »[912]. L'informalité est donc plus contrainte que voulue. Le constat général est que les États soutiennent peu les sociétés nationales d'évaluation. S'ils les acceptent dans la plupart des cas, leur création est très rarement motivée par une impulsion gouvernementale. Les gouvernements restent dans une démarche utilitariste, et non de soutien, vis-à-vis de ces

[909] COLONOMOS, Ariel, dir. *Sociologie des réseaux transnationaux, Communautés, entreprises et individus : lien social et système international*, L'Harmattan, 2001, 300p, p 32.
[910] Nous avons produits ces données à partir d'une enquête réalisée par l'IOCE en 2013.
[911] Représentants de sociétés d'évaluation ayant participé à l'enquête citée ci-dessus.
[912] Entretien avec la coresponsable du réseau d'évaluateurs IOCE, ancienne présidente de l'IPEN, 22 janvier 2014.

réseaux. Plusieurs responsables de sociétés d'évaluation témoignent que si leur gouvernement a reconnu leur réseau d'évaluation, aucune collaboration n'a pour autant été établie[913].

Nos lectures et les entretiens que nous avons menés montrent que les réseaux d'évaluateurs veulent s'institutionnaliser. Les sociétés informelles sont souvent celles qui ont été créées le plus tardivement et nous retrouvons dans le discours de leurs représentants une volonté forte d'obtenir la reconnaissance des gouvernements. Ainsi, un membre de l'association marocaine d'évaluation explique que « *la société civile était faible et éclatée sur le thème de l'évaluation et n'exerçait aucune influence sur les décideurs. C'est alors qu'un petit groupe de consultants et d'universitaires, ayant constaté la demande des donateurs internationaux en matière d'évaluation, a saisi l'opportunité pour se constituer en société informelle. Puis, ils ont voulu passer du stade des discussions informelles à celui d'une organisation formelle. Le travail n'a pas pu aboutir aussi vite que souhaité car il a fallu mener une campagne d'explication auprès des décideurs politiques et surtout auprès des organisations internationales pour qu'ils soutiennent ce projet de formalisation d'une société de suivi-évaluation. En 2004, l'association marocaine d'évaluation a été reconnue* »[914]. Le président de la société sénégalaise d'évaluation considère que son caractère informel « *empêche la coopération avec le gouvernement et son soutien. La transition vers une association formelle est l'un des objectifs prioritaires du groupe à court terme* »[915]. Pour sa part, la co-responsable du réseau d'évaluateurs Evalpartners estime que « *pour assurer la durabilité des réseaux, il faut des professionnels, des personnes dont le travail est de les faire vivre. Il faut un secrétariat et des locaux. Toutefois, dans la majorité des réseaux, il n'y a pas assez de membres pour couvrir ces dépenses* »[916].

Une caractéristique des réseaux nationaux et transnationaux, qui va de pair avec l'informalité, est qu'ils sont fondés sur des rapports horizontaux et non hiérarchiques. C'est pourquoi les responsables des réseaux transnationaux d'évaluation mettent en avant la dimension de partenariat égalitaire entre le Nord et le Sud dans leurs discours, fonctionnement que nous allons remettre en cause.

Selon Ariel Colonomos, le réseau est conçu « comme un lien qui tire sa force de sa faible institutionnalisation, de la promotion de la logique associative sur celle de l'autorité

[913] Représentants de sociétés d'évaluation ayant participé à l'enquête de l'IOCE.
[914] Entretien avec un évaluateur marocain, 17 septembre 2014.
[915] Entretien avec le directeur d'un cabinet d'évaluation au Sénégal, président de l'association sénégalaise d'évaluation, 3 avril 2014.
[916] Entretien avec la coresponsable du réseau d'évaluateurs IOCE, ancienne présidente de l'IPEN, 22 janvier 2014.

hiérarchisée et administrée (…). Le réseau ne suppose pas nécessairement, contrairement à l'institution, un centre hiérarchique et une organisation verticale. Bien au contraire, le réseau obéit à une logique associative et se déploie dans l'horizontalité des rapports sociaux, qui fonde la spécificité de son fonctionnement »[917].

D'autres chercheurs sont plus sceptiques quant à la capacité des réseaux internationaux à rendre plus égalitaires les relations entre les acteurs du système international. Margaret Keck et Kathryn Sikkink estiment par exemple que « pour les acteurs du Tiers monde, les réseaux fournissent des accès, des leviers, de l'information et des moyens financiers qu'ils ne pourraient pas avoir par eux-mêmes. Cela donne l'impression aux acteurs du Nord de travailler avec, et non pas seulement pour, leurs partenaires du Sud. Mais il n'est pas rare de voir se reproduire au sein des réseaux les relations de pouvoirs Nord-Sud qu'ils ont pourtant vocation à dépasser »[918].

Nos recherches nous ont conduite à penser que la création des sociétés d'évaluation, quelle que soit leur échelle d'action, s'est faite en réaction aux rapports hiérarchisés et verticaux qu'entretenaient les organisations d'aide au développement avec les bénéficiaires de l'aide dans le cadre des interventions de renforcement des capacités en évaluation. Les réseaux d'évaluation ont bien vocation à créer de nouveaux espaces d'échange et d'apprentissage mutuel plus égalitaires entre le Nord et le Sud.

Toutefois, nous observons que les rapports de force entre le Nord et le Sud sont toujours présents dans ces réseaux. Premièrement, les sociétés d'évaluation nationales du Nord et du Sud établissent peu de contacts. Il n'y a pas de coopération entre ces deux types d'acteurs, ce qui rend difficile de parler d'un « réseau mondial des évaluateurs ». Les réseaux du Nord sont ancrés dans une perspective purement nationale et leur mandat ne comprend souvent ni le renforcement des capacités des évaluateurs du Sud, ni le soutien aux réseaux du Sud[919]. Les seuls cas de coopération se font de personne à personne en raison de liens amicaux ou professionnels crées lors de rencontres ou de collaborations passées. En revanche, nous notons une collaboration significative Nord-Nord et Sud-Sud entre les sociétés d'évaluation, en termes d'échanges de bonnes pratiques et de connaissances. En conséquence, l'opportunité, ouverte par la constitution de réseaux nationaux d'évaluation au Sud de renforcer les liens entre les

[917] COLONOMOS, Ariel, 2001, *op. cit*, p. 23.
[918] KECK, Margaret, SIKKINK, Kathryn, 1998, *op. cit.* p. 131.
[919] Entretien avec un consultant en évaluation des politiques publiques au Maroc, 17 septembre 2014.

évaluateurs du Nord et ceux du Sud et ainsi de favoriser l'implication de ces derniers dans les évaluations des programmes d'aide au développement[920], nous apparaît limitée.

S'il y a un renforcement des liens, il passe par les réseaux de plus large échelle, de niveau régional (de nombreux évaluateurs du Nord participent aux congrès de l'AFREA par exemple) ou international. Mais même dans ces réseaux, les liens créés restent très superficiels car ils ne se font que ponctuellement lors de séminaires annuels où chacun rencontre de multiples personnes. Le co-responsable du réseau Evalpartners explique que « *c'est pour combler ce manque de dialogue entre évaluateurs du Nord et du Sud que le réseau Evalpartners a été créé en 2012, afin de former un réseau d'évaluateurs réellement transnational* »[921]. Selon nous, la création de réseaux internationaux n'est qu'un point de départ à la formation de liens professionnels entre évaluateurs du Nord et du Sud. De tels liens, forts et pérennes, ne peuvent se nouer que par la pratique partenariale de l'évaluation dans le cadre de programmes de développement. S'il n'existe pas d'opportunités et une volonté partagée pour concrétiser par un partenariat les rencontres permises par les réseaux internationaux, une communauté internationale d'évaluateurs ne pourra pas émerger.

Cette faiblesse des liens entre évaluateurs du Nord et du Sud, la méconnaissance et les *a priori* qui persistent entre les deux groupes malgré la création de réseaux d'évaluateurs s'expriment bien lorsque nous interrogeons les évaluateurs sur ce problème. Nous nous sommes demandé pourquoi les évaluateurs du Nord ont tant de réticences ou de difficultés à réaliser leurs travaux de manière conjointe avec leurs pairs du Sud, question à laquelle ne répondait pas l'enquête de l'IOCE. Nous avons alors interrogé des évaluateurs du Nord et du Sud pour avoir leur opinion sur ce questionnement.

En premier lieu, les évaluateurs du Nord pensent généralement que les compétences des experts du Sud en évaluation sont faibles. Un évaluateur exerçant au sein d'un cabinet d'audit parisien, déclare, par exemple, à propos d'expériences passées : « *la vraie difficulté était d'être associé avec des partenaires qui n'avaient pas l'habitude de mener des entretiens, de conduire des enquêtes et de structurer des rapports d'évaluation* »[922]. Les évaluateurs du Nord témoignent également de la difficulté de constituer des équipes Nord-Sud compte tenu des exigences des appels d'offre des bailleurs de fonds. Un évaluateur dans un cabinet belge

[920] Considérations relevées lors d'un atelier de travail du réseau Evalnet du CAD-OCDE le 30 janvier 2014. Voir liste des évènements observés.
[921] Entretien avec un évaluateur des Nations unies, 6 octobre 2014.
[922] Considérations relevées lors d'un atelier de travail du réseau Evalnet du CAD-OCDE le 30 janvier 2014. Voir liste des évènements observés.

spécialisé en évaluation des politiques publiques, déclare à ce sujet : « *le problème principal est de constituer une équipe mixte qui réponde aux critères des appels d'offres. Ajouter un consultant du Sud peut être très difficile, car les appels d'offres contiennent des contraintes fortes en termes d'expérience et de capacités spécialisées qui sont laborieuses à trouver rapidement au Sud, d'autant plus lorsque l'on n'a pas une bonne connaissance des compétences y existent. C'est un vœu pieu que d'espérer que les consultants locaux participent à l'évaluation dès la réponse aux appels d'offres qui se fait dans l'urgence* »[923]. Ici, nous voyons que la mise en réseau d'évaluateurs du Nord et du Sud est un réel besoin. Toutefois, si cette mise en réseau peut favoriser la constitution d'équipes d'évaluation Nord/Sud, elle ne garantit pas que cette coopération soit égalitaire.

Le faible degré de coopération entre évaluateurs du Nord et du Sud reflète également un manque de confiance mutuelle. Ce point est particulièrement soulevé par les évaluateurs du Sud. Un évaluateur burkinabé déclare qu'« *il y a un problème de reconnaissance au Nord des compétences du Sud. Même lorsque des évaluateurs du Sud sont intégrés aux évaluations, leur expertise est peu valorisée dans les rapports finaux, elle est même parfois reléguée en annexe* »[924].

Ce manque de confiance s'exprime aussi dans la crainte qu'ont les évaluateurs du Nord d'un manque d'objectivité des experts du Sud. L'un d'entre eux souligne que « *le consultant manque souvent d'indépendance, car il appartient à l'administration évaluée, ou bien rêve de la rejoindre. Les évaluateurs du Sud sont trop impliqués dans les réseaux de la fonction publique de leurs pays pour être réellement indépendants* »[925]. Nous voyons donc que bon nombre de présupposés subsistent entre les évaluateurs du Nord et du Sud, que la création de réseaux d'évaluateurs n'a, pour l'instant, pas permis de dépasser.

Une autre caractéristique des réseaux d'évaluateurs est qu'ils fonctionnent sur la base du bénévolat et avec peu de moyens. Leur bon fonctionnement repose avant tout dans les individus qui les composent. Les réseaux se fondent et se développent grâce à l'implication continue de personnes engagées dans la promotion de l'évaluation dans leur pays ou au niveau international. La force d'un réseau réside essentiellement dans le nombre de membres qu'il réunit, notamment de membres actifs. C'est pourquoi nous voyons dans le très faible nombre de membres des sociétés nationales d'évaluation en Afrique un facteur qui remet en cause leur durabilité. Plus

[923] Considérations relevées lors d'un atelier de travail du réseau Evalnet du CAD-OCDE le 30 janvier 2014. Voir liste des évènements observés.
[924] *Ibid.*
[925] *Ibid.*

le nombre de membres est faible, moins le réseau va pouvoir générer de fonds et assurer sa pérennité. Plus le réseau repose sur un petit socle de membres actifs, plus le risque est grand qu'il périclite, dans la mesure où il s'agit de bénévoles qui ont d'autres priorités. Ainsi s'intéresser aux réseaux, c'est réévaluer le rôle de l'acteur dans la diffusion d'une culture de l'évaluation, processus qui est souvent étudié de manière très institutionnelle par la littérature qui se concentre sur les processus de transferts de normes d'évaluation d'organisation à organisation.

Cette centralité des individus dans l'activité des réseaux est à la fois une force et une faiblesse. Selon la directrice de l'IPEN et cofondatrice du réseau Evalpartners, « *la durabilité des sociétés d'évaluation dépend des personnes qui les fondent. Dans toutes les sociétés, il y a eu des hauts et des bas, beaucoup ont failli s'effondrer ou se sont arrêtées et ont repris par la suite. Lorsque vous dirigez une société professionnelle d'évaluation, votre budget, ce n'est pas de l'argent, c'est de la sueur. La clef de la réussite est d'avoir des personnes qui ont des idées, qui sont prêtes à s'investir sur le long terme et régulièrement. Dans beaucoup de pays, j'ai vu ce genre de personnes. Lorsqu'il y a un leader fort, généralement cela fonctionne bien* »[926]. D'autres interviewés voient davantage la centralité des individus dans les réseaux comme un handicap à leur bon fonctionnement. C'est le cas, par exemple, d'une chargée de projet au sein du F3E, pour qui « *le degré d'efficacité et d'activité des réseaux d'évaluation est beaucoup une question de personnes. Dans la mesure où les réseaux ont généralement peu de moyens, les membres sont souvent des bénévoles qui ont une forte tendance à l'immobilisme. Si l'on prend l'exemple du Réseau francophone de l'évaluation, son fonctionnement est presque une question d'amitié et d'inimitié entre les membres. Comme les réseaux sont relativement petits, ils se retrouvent en perte de vitesse dès que les personnes leaders s'impliquent moins. Cette dépendance à un petit nombre de personnes entraîne facilement des situations de blocage* »[927]. La prolifération des réseaux d'évaluation au début des années 2000 reflète un contexte international où, de manière générale, l'individu en tant qu'acteur politique prend de l'ampleur, notamment grâce au développement des nouveaux médias (réseaux sociaux, blogs, etc.) qui permettent la mise en réseau d'individus éloignés dans l'espace national ou international. Ce constat nous amène à nous questionner sur l'importance des « nouvelles technologies » dans le fonctionnement des sociétés d'évaluation.

[926] Entretien avec la coresponsable du réseau d'évaluateurs IOCE, , 22 janvier 2014.
[927] Entretien avec une chargée de projet en évaluation au F3E, 19 avril 2012.

- **L'importance des nouvelles technologies dans la socialisation des évaluateurs**

D'après nos recherches, le rôle des nouvelles technologies est très important dans le développement des réseaux d'évaluation, notamment au Sud. La quasi intégralité des échanges entre les membres passe par Internet. Les réseaux les plus avancés ont un site Internet et/ou une page Facebook et la plupart communiquent *via* des listes de diffusion par courriers électroniques. Les échanges numériques constituent la vie quotidienne des réseaux pour pérenniser leur action au-delà des réunions ponctuelles qu'ils organisent.

Dans la majorité des cas, la seule activité des réseaux, au début de leur existence, consiste en des échanges électroniques entre leurs quelques membres qui peuvent être éloignés géographiquement. Certains réseaux ont donc d'abord une existence virtuelle, avant de s'ancrer dans le réel par l'organisation d'évènements réunissant physiquement leurs membres. Nous pouvons donc penser que, sans ces outils informatiques, la pérennité et le développement des réseaux d'évaluation seraient compromis. Cette importance d'Internet dans le fonctionnement des réseaux transnationaux est largement soulignée par la littérature existante sur le sujet. Par exemple, Margaret Keck et Kathryn Sikkink affirment que « les échanges d'informations lient les membres des réseaux et sont essentiels à leur efficacité. Beaucoup d'échanges d'information sont informels, par téléphone, e-mails, fax et la diffusion de bulletins d'information. Les réseaux fournissent une information qui ne serait pas disponible autrement à leurs membres, provenant de sources qui ne seraient pas entendues sans leur intermédiaire et la font circuler auprès de publics qui peuvent être éloignés géographiquement »[928]. Ces échanges informels permettent aux membres des réseaux d'échanger rapidement des informations alternatives à celles produites et diffusées par les instances gouvernementales.

Au niveau des sociétés régionales et internationales d'évaluation, l'utilisation d'Internet a pour principal bénéfice de démocratiser l'accès aux activités de renforcement des capacités en évaluation qu'elles organisent. Les réseaux retransmettent sur Internet les ateliers de travail, séminaires et conférences qu'ils organisent. Ils élaborent également des cours spécialement conçus pour être diffusés par ce vecteur. Ainsi, l'influence des réseaux est plus large que si ces activités étaient seulement organisées de manière physique.

Toutefois, l'importance des outils numériques pour la socialisation des membres du réseau pose la question de l'accès au numérique dans les pays en développement. En effet, la mise en réseau apparaît difficile pour les personnes qui ne disposent pas d'un accès à Internet.

[928] KECK, Margaret, SIKKINK, Kathryn, 1998, *op. cit.* p. 21.

Bien que cet accès s'élargisse dans les pays du Sud, de nombreuses zones rurales en restent encore privées. Cependant, les sociétés d'évaluation visent à constituer une communauté d'élites : des représentants gouvernementaux, des responsables d'OSC, des professionnels de l'évaluation ou des chercheurs, qui ont, notamment de par leur profession, accès à Internet. Le problème de la fracture numérique ne se pose donc pas de manière significative pour la survie des réseaux d'évaluateurs. Toutefois, il est réel en ce qui concerne la diffusion à la société civile des travaux produits par ces réseaux, afin de sensibiliser la population à l'évaluation des politiques publiques et de créer ainsi une demande populaire pour une hausse de la transparence et de la redevabilité dans les politiques publiques.

Enfin, nous avons étudié de manière plus approfondie le type d'activités que les sociétés d'évaluation mettent en place[929]. Les sociétés au Sud organisent surtout des formations pour leurs membres sous la forme d'ateliers de travail ou de journées d'études. Une autre activité récurrente est l'organisation de « journées nationales de l'évaluation » annuelles, afin de donner une visibilité au réseau dans l'espace public national et de susciter l'intérêt de potentiels nouveaux membres. Ces conférences sont également l'occasion pour les acteurs de l'évaluation (opérateurs, commanditaires, utilisateurs) de se connaître et d'échanger. C'est souvent grâce à ces évènements que les sociétés nationales d'évaluation se créent et qu'elles gagnent de nouveaux membres. De plus, la moitié des sociétés nationales d'évaluation recensées par l'IOCE financent la participation de leurs membres aux conférences organisées par les sociétés régionales d'évaluation, afin d'échanger des connaissances et des expériences dans un cercle plus large. Toutefois, rares sont celles qui soutiennent la participation de leurs membres aux séminaires organisés par les sociétés internationales d'évaluation. Ce sont davantage ces dernières qui financent la venue des intervenants qu'elles ciblent dans les sociétés nationales d'évaluation comme étant des intervenants potentiellement intéressants.

Nous constatons également que rares sont les sociétés d'évaluation qui agissent auprès des gouvernements pour les conseillers dans leurs prises de décision concernant le suivi et l'évaluation des politiques publiques. Pourtant, comme nous l'avons vu ci-dessus, influencer la gouvernance de l'État est un objectif affiché de ces réseaux d'acteurs. Les raisons de la faiblesse de ce pan d'actions seront étudiées dans la suite de ce chapitre. Enfin, rares sont les sociétés nationales du Sud à développer des actions concrètes pour développer la recherche et les publications sur l'évaluation.

[929] Cette étude sur les activités menées par les réseaux d'évaluateurs a été réalisée à partir de l'enquête réalisée par l'IOCE en 2013, portant sur l'état actuel de ces réseaux dans le monde.

En ce qui concerne les sociétés régionales d'évaluation, leurs actions principales sont l'organisation de conférences à destination de tous les évaluateurs et membres de sociétés d'évaluation de leur zone géographique d'influence. Elles apportent également un soutien aux sociétés nationales d'évaluation, aussi bien financier qu'en expertise. De plus, elles développent des activités de formation et de recherche sur l'évaluation. Une activité qui les distingue des sociétés nationales d'évaluation est la production de bonnes pratiques. La production de normes et de bonnes pratiques est l'activité des réseaux la plus fréquemment citée par la littérature théorique, par exemple dans les travaux de Margaret Keck et Kathryn Sikkink : « les réseaux promeuvent la convergence et l'harmonisation des normes aux niveaux régional et international. Ils encouragent également la mise en œuvre des normes en exerçant une pression sur des acteurs ciblés pour qu'ils adoptent de nouvelles politiques »[930].

Une ancienne présidente de l'AFREA, raconte que ce réseau d'évaluateurs est né du projet de développer un ensemble de « bonnes pratiques » d'évaluation communes au continent africain : « *en 1998 un atelier de travail régional a été organisé à Nairobi par l'UNICEF. Une des sessions fut consacrée au thème : les normes des programmes d'évaluation aux États-Unis sont-elles appropriées aux cultures africaines ? Cet atelier de travail a été l'élément déclencheur de la création de l'AFREA. Il était animé par Mahesh Patel qui est ensuite devenu le premier président de ce réseau* »[931]. À la suite de cette session, un travail plus approfondi a été mené par les participants pour réfléchir à la possibilité du transfert de ces bonnes pratiques américaines au contexte du continent africain. Notre interviewée raconte que les résultats de ce travail furent présentés à la conférence inaugurale de l'AFREA en 1999, à laquelle ont participé plus de 300 évaluateurs africains. Mahesh Patel recommanda que les bonnes pratiques africaines de l'évaluation soient revues par les réseaux nationaux d'évaluation et testées sur le terrain. Ce test empirique a eu lieu dans deux pays, la Zambie et le Kenya et dans le cadre d'une évaluation multi-pays. Il a été concluant et les bonnes pratiques de l'AFREA sont diffusées à de nombreux gouvernements en Afrique. La légende veut que ces bonnes pratiques africaines aient influencé les standards sur l'évaluation élaborés en 2005 par le réseau d'évaluateurs des Nations unies (l'UNEG). Elles sont fréquemment révisées lors des conférences annuelles de l'AFREA[932].

Ce récit laisse transparaître un exemple des jeux d'acteurs qui se construisent autour des « bonnes pratiques » de l'évaluation des programmes de développement. En effet, l'année

[930] KECK, SIKKINK, 1998, *op. cit.* p. 36.
[931] Entretien avec une ancienne présidente de l'AFREA, 22 mai 2012.
[932] Entretien avec une ancienne présidente de l'AFREA, 22 mai 2012.

même (en 2005) où l'OCDE a produit ses propres bonnes pratiques sur l'efficacité de l'aide, décriées pour être « occidentalo-centrées », les Nations unies ont créé, de leur côté, un ensemble de bonnes pratiques affiché comme étant directement inspiré des standards élaborés par les pays africains et donc centré sur les besoins des pays bénéficiaires. Les Nations unies se sont donc placées en alternative au CAD-OCDE en proposant une approche *bottom-up* d'élaboration de standards internationaux de l'évaluation, se voulant en rupture avec l'approche *top-down* et en cercle fermé du CAD. Malgré cette instrumentalisation de l'AFREA par les Nations unies, sur laquelle nous reviendrons plus tard, le processus d'élaboration de bonnes pratiques dans le cadre d'une société d'évaluation du Sud est novateur. Il a permis la participation, au premier plan, d'évaluateurs du Sud qui ont pu dialoguer avec les représentants des organisations internationales qui supervisaient l'élaboration des bonnes pratiques.

En plus de ces activités de formation et d'organisation d'évènements sur le thème de l'évaluation, les réseaux internationaux d'évaluation ont récemment mis en place des blogs consacrés au *storytelling* sur l'évaluation. Des appels à contribution internationaux sont en cours de diffusion pour que les évaluateurs partagent des récits présentant des cas d'évaluations ayant mené à « changer des vies »[933]. Nous y voyons une tentative de plaidoyer visant à persuader « l'opinion publique internationale » que l'évaluation est utile et peut faire une différence dans les processus de développement des pays du Sud. Ce recours au *storytelling*, comme outil de persuasion est caractéristique de l'action des réseaux transnationaux. Nous retrouvons beaucoup de référence à ce mode de communication dans la littérature théorique sur les réseaux, par exemple dans les travaux de Keck et Sikkink : « les flux d'information dans les réseaux ne fournissent pas seulement des faits, mais aussi des *témoignages* – des histoires racontées par des personnes dont les vies ont été affectées. Le processus selon lequel les histoires sont découvertes et présentées suit plusieurs étapes. Premièrement, les acteurs transnationaux identifient les types de récits utiles à leur plaidoyer et lancent des campagnes pour trouver les personnes à même de les raconter. Ensuite, ces histoires sont filtrées et adaptées par les membres du réseau. Il y a fréquemment un écart énorme entre l'histoire originelle et sa forme adaptée. En d'autres termes, les personnes locales perdent souvent le contrôle sur le contenu de leurs récits au cours de la « campagne » transnationale »[934].

Les blogs étant encore en construction en 2014, il est, toutefois encore impossible de juger de l'ampleur de l'écart qu'il y aura entre les témoignages et la version finale des histoires

[933] Le projet de blog et l'appel à consultation sont consultables sur le site du réseau d'évaluateurs Evalpartner : https://evaluationstories.wordpress.com/about/, consulté le 23 avril 2013.
[934] KECK, Margaret, SIKKINK, Kathryn, 1998, *op. cit.* p. 200.

présentées au public. Le réseau Evalpartners a également lancé une campagne internationale pour faire valoir l'évaluation en proclamant l'année 2015 « année internationale de l'évaluation ». Les responsables de ce réseau y voient l'opportunité de catalyser les intérêts sur l'évaluation et de dynamiser un ensemble de réseaux dont beaucoup restent peu actifs[935].

C. Une faible influence sur les gouvernements du Sud

Après avoir étudié le processus de formation des sociétés nationales, régionales et internationales d'évaluation, ainsi que leur mode de fonctionnement, il nous a paru intéressant de nous questionner sur l'efficacité des réseaux d'évaluateurs. De quelle manière transforment-t-ils le système international dans lequel ils sont intégrés ?

- **Des réseaux d'évaluateurs relais de l'entreprise de normalisation des organisations multilatérales**

Les chercheurs travaillant sur la problématique des réseaux mesurent l'efficacité de ces derniers par rapport à l'influence qu'ils ont sur les acteurs du système international, notamment sur les États et les organisations internationales. Ils sont jugés influents si leur expertise et leur plaidoyer parviennent à faire changer les positionnements et les comportements de ces acteurs sur des enjeux ciblés. Autrement dit, ils ont rempli leur rôle lorsque l'enjeu pour lequel ils se mobilisent rentre sur l'agenda politique national ou international.

Keck et Sikkink retiennent cinq dimensions pour évaluer l'efficacité d'un réseau transnational : (1) la construction d'un enjeu et sa mise sur agenda politique ; (2) son influence sur les positions discursives des États et des organisations régionales et internationales ; (3) son influence sur les procédures institutionnelles de ces acteurs ; (4) son influence sur les changements de politiques de ces acteurs ; (5) son influence sur le comportement des États[936]. Ce modèle se focalise sur l'efficacité externe des réseaux, c'est-à-dire sur leur influence sur d'autres acteurs de la scène internationale.

Pour ce qui est des réseaux transnationaux d'évaluation, qui ne sont pas, comme c'est le cas dans l'analyse de Keck et Sikkink, à proprement parler des « réseaux de plaidoyer », nous estimons qu'il n'est pas approprié de juger de leur efficacité seulement sur le plan de leur

[935] KOSHELEVA, Natalia, SEGONE, Marco, 2013, *art. cit.*
[936] KECK, Margaret, SIKKINK, Kathryn, 1998, *op. cit.*

influence externe. D'autres facteurs sont à prendre en compte, notamment leur aptitude à renforcer les capacités en évaluation de leurs membres et à construire une culture commune de l'évaluation parmi des évaluateurs opérant dans des contextes divers.

Les sociétés d'évaluation ont réussi à s'instituer comme un acteur indispensable du renforcement des capacités en évaluation, proposant une forme d'apprentissage alternative aux organisations internationales. Leur efficacité en la matière est reconnue par les acteurs du développement. Un évaluateur travaillant sur les enjeux de renforcement des capacités en évaluation, déclare par exemple que « *les sociétés d'évaluation sont une partie du renforcement des capacités en évaluation. Une association dynamique permet de diffuser les valeurs liées à l'évaluation de façon large. Elle permet de diffuser les normes de l'évaluation, de créer des garde-fous, de la régulation. Elles sont indispensables, mais il faut avant tout des mécanismes gouvernementaux qui permettent la pratique de l'évaluation au Sud* »[937]. La co-responsable d'Evalpartners a aussi une opinion positive de l'efficacité des sociétés d'évaluation en termes de renforcement des capacités : « *les sociétés d'évaluation sont efficaces en ce qui concerne le renforcement des capacités en évaluation. Elles organisent des conférences pour que les évaluateurs se retrouvent et partagent leurs expériences, ce qui crée de l'émulation. Les membres des réseaux du Sud peuvent partager des expériences qui sont pertinentes par rapport à leur contexte national, ce qui n'est pas forcément le cas lors d'échanges avec des évaluateurs du Nord et des organisations internationales. Le dialogue entre pairs permet une meilleure diffusion des bonnes pratiques qu'un dialogue de type dominant-dominé* »[938].

D'aucuns estiment que la première réussite des réseaux d'évaluation serait donc d'avoir créé une communauté internationale d'évaluateurs partageant des compétences et des valeurs communes[939]. Nous tenons à nuancer cette affirmation. Bien que les réseaux permettent une diffusion horizontale, par les pairs, de la culture de l'évaluation qui est indispensable, il ne faut pas surestimer leur efficacité potentielle en matière de renforcement des capacités en évaluation. En effet, les évaluateurs auront beau échanger leurs connaissances et leurs expériences, s'il n'y a pas de demande suffisante d'évaluation au niveau national, permettant aux évaluateurs de se former sur le terrain en pratiquant l'évaluation, l'activité des réseaux ne servira pas à grand-chose.

[937] Entretien avec le directeur d'un cabinet d'évaluation, consultant en évaluation auprès du CAD-OCDE, 20 novembre 2013.
[938] Entretien avec la coresponsable du réseau d'évaluateurs IOCE, 22 janvier 2014.
[939] Entretien avec le directeur d'un cabinet d'évaluation des politiques publiques, consultant en évaluation auprès du CAD-OCDE, 20 novembre 2013.

À ce stade, il faut rappeler que la plupart des réseaux d'évaluation sont des constructions sociales très jeunes, dont un tiers a, au plus, cinq années d'existence. En outre, leur création ayant été, le plus souvent, largement motivée par les incitations des organisations internationales et concrétisée grâce au regroupement de quelques individus, il leur a fallu, en premier lieu, construire une communauté de membres. En effet, la population des évaluateurs est très éparpillée et peu visible au niveau national et très fragmentée au niveau international. C'est pourquoi nous estimons que le fait d'avoir réussi à créer des « communautés de pratique » au niveau national ou régional peut déjà être vu comme une réussite. Cependant, ce processus est loin d'être achevé. En effet, nous avons vu que le dialogue entre les évaluateurs du Nord et du Sud reste faible. La question de l'existence d'un réel réseau transnational d'évaluateurs se pose donc. Nous allons démontrer qu'il existerait plutôt un réseau du Nord et un réseau du Sud, qui resteraient peu intégrés. Les réseaux internationaux, tels qu'Evalnet ou l'IOCE, tendent à dépasser cette dichotomie et à créer un lien entre les évaluateurs du Nord et du Sud. Toutefois, ces réseaux internationaux sont le fruit des organismes multilatéraux et les échanges entre les évaluateurs se font donc dans le cadre d'une relation entre donateurs et bénéficiaires peu favorable à un échange égalitaire.

- **Une influence politique très limitée**

Au-delà du renforcement des capacités de leurs membres en matière d'évaluation, les sociétés d'évaluation se donnent comme objectif d'influencer les politiques des États du Sud vers une pratique plus intensive de l'évaluation. Elles ont aussi pour objectif de créer une demande et une offre en évaluation au Sud pour rendre les évaluations des programmes d'aide plus inclusives des acteurs du Sud et des enjeux politiques qui les préoccupent. Cela nous amène à nous questionner sur l'influence politique qu'ont les réseaux d'évaluateurs sur les organismes d'aide et les gouvernements des États du Sud.

Les évaluateurs, du Nord comme du Sud, défendent l'idée que les réseaux régionaux, notamment du Sud, influencent la politique d'évaluation des organisations internationales. Le *storytelling* le plus couramment cité pour prouver cette affirmation est l'histoire évoquée ci-dessus, selon laquelle les Nations unies auraient été influencées par l'AFREA dans l'élaboration de leurs standards d'évaluation[940]. Notre analyse nous amène à remettre en cause cette influence présumée. En effet, l'AFREA a été mise en place grâce à l'UNICEF de qui elle reçoit toujours

[940] Ce récit a été identifié dans plusieurs sources : entretien avec une ancienne présidente de l'AFREA, 22 mai 2012 ; RUGH, Jim, SEGONE, Marco, 2013, *op. cit.*

un soutien important. Les fondateurs et les dirigeants de l'AFREA ont tous participé à des activités de renforcement des capacités organisées par les Nations unies et ont été marqués par leurs discours sur l'évaluation. Certains sont également membres des comités de direction de sociétés transnationales d'évaluation, également sous l'influence directe des organisations multilatérales, en premier lieu desquelles les Nations unies. De plus, les membres les plus actifs de l'AFREA ont, pour la plupart, travaillé par le passé auprès de donateurs bilatéraux ou multilatéraux dans la mise en œuvre d'évaluations de programmes d'aide. Ces acteurs ont donc été nourris, pendant de longues années, des bonnes pratiques des donateurs et de leurs approches de l'évaluation.

Parler d'une influence des membres de l'AFREA sur les pratiques d'évaluation des Nations unies apparaît donc trop ambitieux. L'influence normative en matière d'évaluation se fait davantage des organisations internationales vers les sociétés d'évaluation que dans le sens inverse. Les réseaux d'évaluation restent le fruit et le nouvel instrument des organisations internationales pour diffuser leurs « bonnes pratiques » de l'évaluation à travers le monde. C'est uniquement la forme de l'approche qui a changé, passant d'un schéma *top down* à un schéma *bottom to bottom*, les dirigeants des sociétés d'évaluation se faisant, consciemment ou inconsciemment, les porte-paroles des évaluateurs des organisations internationales. Nous voyons transparaître ici une illusion courante dans les relations internationales actuelles qui consiste à croire qu'à partir du moment où des relations se construisent sur le mode du réseau, elles sont égalitaires. Or, derrière la désignation de personnes de la société civile du Sud à la tête des sociétés d'évaluation se cache une influence forte des organisations internationales qui parrainent ces sociétés. Par exemple, une ancienne présidente de l'IPEN déclare « *s'être fortement appuyée sur l'expérience et l'expertise de personnes telles que Marco Segone ou Robert Picciottio, respectivement anciens dirigeants des départements d'évaluation de l'UNICEF et de la Banque mondiale* »[941].

Les seuls réseaux d'évaluation qui nous apparaissent pouvoir influencer la politique des organismes d'aide au développement sont les réseaux institutionnalisés au sein des organisations multilatérales, par exemple l'UNEG, qui regroupe les unités d'évaluation des agences des Nations unies, Evalnet, qui représente les évaluateurs des agences bilatérales d'aide au développement des États membres du CAD-OCDE, ou encore *l'Evaluation Coordination Group*, qui rassemble les évaluateurs des banques multilatérales de développement. Les organisations multilatérales ont créé ces réseaux car elles avaient besoin de coordonner et

[941] Entretien avec la coresponsable du réseau d'évaluateurs IOCE, 22 janvier 2014.

d'harmoniser les pratiques d'évaluation de leurs diverses entités. Elles avaient également besoin de capitaliser les connaissances issues des évaluations. Elles étaient donc dans un contexte d'incertitude et souhaitaient créer une communauté épistémique d'évaluateurs en leur sein. L'action de ces réseaux dans les organisations internationales permet d'assurer que les enjeux de l'évaluation (avancées méthodologiques, renforcement des capacités en évaluation, production de normes, etc.) restent sur l'agenda politique. Si nous nous référons au cas du réseau Evalnet, nous constatons qu'il a une influence cognitive significative sur le travail du CAD-OCDE, notamment sur la manière dont cette structure aborde les enjeux de l'évaluation en termes de genre, d'États fragiles, ou encore de bonne gouvernance. En effet, le réseau a créé des groupes de travail spécifiques sur ces thèmes, chargés d'influencer la politique du CAD-OCDE et de ses membres, essentiellement *via* la publication de guides de bonnes pratiques et l'organisation de séminaires. Le réseau a également collaboré avec d'autres communautés épistémiques du CAD-OCDE, afin de mettre certains enjeux en haut de l'agenda politique. Par exemple, la collaboration entre Evalnet et le réseau sur le genre depuis les années 1990 a mené à un engagement des membres du CAD-OCDE à systématiquement prendre en compte la question du genre dans l'élaboration, la mise en œuvre et l'évaluation de leurs programmes de développement. Des bonnes pratiques en termes de gestion des problématiques de genre ont été mises en place par le CAD-OCDE. Leur application par les donateurs est régulièrement évaluée par le réseau Evalnet pour maintenir l'attention de ces derniers sur la problématique du genre[942].

Les réseaux placés au sein d'organisations multilatérales puissantes ont donc une influence politique significative dans ces structures. Dans le cas du CAD-OCDE, l'influence est double puisqu'elle porte également sur les États membres de l'organisation. Elle est due au fait que les réseaux sont institutionnalisés au sein des organisations multilatérales. Ces réseaux ne sont toutefois pas comparables aux sociétés nationales et régionales d'évaluation, puisqu'ils ne sont pas composés à proprement parler de membres de la société civile, mais de fonctionnaires internationaux ayant pour la plupart eu une longue carrière dans l'administration de leur État d'origine et dans les organisations internationales. Les contacts qu'ils ont établis dans ces milieux accroissent donc leur influence politique.

Enfin, l'influence des réseaux d'évaluation se mesure par rapport à leurs effets sur la gouvernance des politiques publiques des États. La question est de savoir si l'action des réseaux d'évaluation entraîne une plus grande mise en œuvre et utilisation des évaluations par les gouvernements ciblés. Comme nous l'avons vu plus haut, cet objectif est bien présent dans le

[942] CAD-OCDE, 2013, *op. cit.*

mandat des sociétés d'évaluation du Sud comme du Nord, qui se donnent la mission d'améliorer la gouvernance des programmes de développement, ainsi que la transparence et la redevabilité de la gestion des politiques publiques des États. Selon nous, l'influence politique et normative des réseaux d'évaluation ne peut pas être jugée en dehors de tout contexte. En effet, comme en témoigne une ancienne présidente de l'IPEN, « *l'influence qu'ont les sociétés d'évaluation sur les politiques et la société civile de leurs pays dépend du contexte national. C'est souvent une question d'opportunités* »[943]. Nos recherches ont montré que les sociétés d'évaluation qui ont le plus d'influence sur les politiques d'un ou de plusieurs gouvernements sont celles qui interviennent dans un contexte déjà favorable à l'évaluation. C'est le cas, par exemple, au Maroc, où plusieurs facteurs se sont combinés pour permettre une influence de la société nationale d'évaluation. Il y avait une demande en matière d'institutionnalisation de la fonction d'évaluation par le gouvernement, des liens personnels entre le responsable de la société d'évaluation et des membres du gouvernement, ainsi qu'un nombre conséquent d'évaluateurs dans le pays. Ainsi, la société nationale d'évaluation fut appelée à aider le gouvernement à rédiger la nouvelle constitution de juillet 2011 dans laquelle la fonction d'évaluation des politiques publiques a été intégrée[944]. Cet environnement favorable et les liens entre la société d'évaluation et le gouvernement ont permis à ses membres de faire un plaidoyer efficace auprès du gouvernement pour qu'il améliore ses pratiques en matière d'évaluation et qu'il l'institutionnalise en tant que principe constitutionnel. Un autre exemple de société d'évaluation ayant influencé la politique d'un gouvernement est la société sud-africaine d'évaluation. Le président sud-africain qui était proche de plusieurs membres de la société nationale d'évaluation leur a demandé conseil lorsqu'il décida de créer une politique nationale d'évaluation en 2011[945].

De manière générale, lorsqu'une politique nationale d'évaluation a été élaborée par des gouvernements ou que le principe d'évaluation a été intégré à la constitution d'un État, les sociétés d'évaluation ont été utilisées pour assister les gouvernements dans cette tâche. Cependant, les exemples restent encore rares aujourd'hui, notamment en raison de la jeunesse de la majorité des réseaux d'évaluation au Sud. Selon la co-responsable d'Evalpartners, « *il est très probable que ce type d'exemples se multiplie dans les années à venir, surtout en Asie du Sud. Les sociétés d'évaluation et les parlementaires sont en train de se rapprocher. On le voit dans les conférences régionales sur l'évaluation des programmes de développement* »[946]. Cette

[943] Entretien avec la coresponsable du réseau d'évaluateurs IOCE, 22 janvier 2014.
[944] Entretien avec un consultant en évaluation des politiques publiques au Maroc,17 septembre 2014.
[945] Exemple tiré de l'enquête réalisée par l'IOCE en 2013 sur l'état des sociétés d'évaluation dans le monde.
[946] Entretien avec la coresponsable du réseau d'évaluateurs IOCE, 22 janvier 2014.

vision du futur reste, toutefois, purement hypothétique, dans la mesure où ces cas de sociétés nationales d'évaluation influentes dans les sphères politiques restent très rares au Sud.

En effet, les responsables des sociétés d'évaluation du Sud, même les plus puissantes, témoignent le plus souvent de leur incapacité à avoir une influence politique. Par exemple, une ancienne présidente de l'AFREA, déclare que « *depuis sa création, en 1999, l'AFREA n'a directement influencé aucun gouvernement du continent en termes de politiques nationales d'évaluation, bien que son soutien aux sociétés nationales d'évaluation contribue à développer un environnement favorable à la diffusion d'une culture de l'évaluation dans les États africains* »[947]. Nous notons donc un écart entre l'impact normatif fort qu'aurait l'AFREA sur les Nations unies et sa faible influence sur les États africains, ce qui tend à confirmer notre argument selon lequel la prétendue influence de l'AFREA sur les nouvelles normes d'évaluation des Nations unies relève plus du *storytelling* (conte) que de la réalité. Cette incapacité à influencer la politique des gouvernements est soulignée par de nombreux représentants de sociétés nationales d'évaluation, par exemple en Mauritanie, en Côte d'Ivoire, ou encore au Chili[948]. Des facteurs structurels peuvent expliquer cette difficulté : une culture des politiques publiques peu transparente où la corruption est importante, un contexte qui donne peu de légitimité aux organisations de la société civile, un faible intérêt des pouvoirs publics pour l'évaluation, etc.

Au-delà de ces facteurs externes expliquant la faible influence politique des sociétés d'évaluation, nous trouvons des facteurs internes à ces sociétés.

Tout d'abord, d'un point de vue organisationnel, l'influence des réseaux augmenterait avec leur densité : « les réseaux sont plus efficaces quand ils sont denses, avec de nombreux acteurs, des connections fortes entre les groupes du réseau et des flux d'information crédibles »[949]. Or, nous avons constaté ci-dessus que les liens horizontaux entre sociétés nationales du Nord et du Sud sont faibles, ce qui est aussi vrai pour les liens entre sociétés régionales. Il existe des liens entre les sociétés nationales du Sud, mais ils restent épisodiques. Ce sont surtout les liens verticaux, entre sociétés nationales et régionales du Sud qui sont forts.

Une difficulté supplémentaire est l'incapacité de beaucoup de sociétés nationales du Sud à rassembler assez de membres pour assurer leur fonctionnement. Nous estimons que la hausse du nombre de membres potentiels dépend du développement d'une demande d'évaluation

[947] Entretien avec une ancienne présidente de l'AFREA, 22 mai 2012.
[948] Constats tirés de l'enquête réalisée par l'IOCE en 2013 sur l'état des sociétés d'évaluation dans le monde.
[949] HAAS, Peter, 1992, *art. cit,* p. 14.

gouvernementale et de la société civile. Or, c'est justement le rôle des sociétés d'évaluation de stimuler une demande en évaluation au sein des gouvernements et de la société civile, les organisations internationales s'étant déchargées de cette mission sur elles. Les sociétés d'évaluation se retrouvent alors dans une position difficile où elles ont besoin d'une hausse de la demande en évaluation pour se pérenniser, hausse qu'ils ont à charge de créer.

Un autre aspect conditionnant l'influence des réseaux transnationaux est qu'ils doivent pouvoir s'appuyer sur des membres ayant des connexions avec les gouvernements[950]. Effectivement, nous avons vu, avec le cas du Maroc et de l'Afrique du Sud, qu'un des facteurs les plus déterminants dans le poids qu'ont les sociétés d'évaluation sur les politiques gouvernementales est l'existence de personnes influentes en leur sein, qui ont des relations de confiance avec les gouvernements. Cependant, dans la plupart des cas, les sociétés d'évaluation du Sud ne comprennent pas de tels membres. Leur travail de plaidoyer et de persuasion auprès des gouvernements pour faire valoir leur expertise va, en conséquence, durer plus longtemps, souvent de nombreuses années. Pour compenser ce manque, les sociétés d'évaluation nationales vont, une fois encore, chercher le soutien des organisations internationales qui peuvent avoir les liens qui leur manquent avec les gouvernements.

Ensuite, pour être influents, il faut que les réseaux d'évaluateurs aient la capacité de mobiliser rapidement des informations crédibles et utiles pour les décideurs politiques. Peter Haas analyse la montée en puissance des réseaux en ces termes : « alors qu'une demande d'information émerge, des réseaux ou communautés de spécialistes capables de produire et de diffuser cette information se forment et prolifèrent. Les membres de ces réseaux deviennent influents à mesure que les décideurs politiques sollicitent leur information et leur délèguent des responsabilités »[951]. Or, dans les pays du Sud, la création des sociétés d'évaluation n'a pas répondu à une demande d'information des gouvernements, mais à une demande des organisations internationales qui y voient le moyen de diffuser une culture de l'évaluation dans les États bénéficiaires de l'aide, tout en se dégageant de cette responsabilité. Les sociétés d'évaluation n'ont donc pas été créées pour répondre à une demande nationale, mais pour la créer. Sur ce point, elles diffèrent des réseaux transnationaux décrits habituellement dans la littérature académique.

Les réseaux transnationaux doivent également avoir la capacité de mettre en place des actions ou des récits symboliques pour donner un sens aux enjeux qu'ils défendent auprès

[950] KECK, Margaret, SIKKINK, Katrhyn, 1998, *op. cit.*
[951] HAAS, Peter, 1997, *op. cit*, pp. 4-5.

d'audiences diversifiées et qui peuvent être dispersées géographiquement. À ce jour, comme nous l'avons vu, seuls les réseaux transnationaux d'évaluation mettent en place des stratégies d'influence fondées sur le *storytelling*. Ces initiatives n'ayant pas encore abouti, il est impossible de statuer sur l'influence politique qu'elles peuvent générer. Toutefois, au Mexique, le *storytelling* créé autour de l'expérience de l'évaluation expérimentale Progresa a permis de renforcer la demande d'évaluation du gouvernement[952]. Nous pouvons donc penser que la communication à partir d'expériences positives d'évaluation au niveau local pourrait conduire à persuader les gouvernements de l'utilité de développer davantage la pratique et l'utilisation des évaluations.

Enfin, toujours selon la théorie des réseaux transnationaux, pour avoir une influence politique, les réseaux doivent être capables de contraindre les acteurs puissants à agir selon les principes qu'ils défendent. Les réseaux nationaux d'évaluation du Sud ne développent pas d'activités normatives, excepté les rares qui ont été appelés à le faire par des gouvernements souhaitant développer des politiques d'évaluation. Les mandats des sociétés régionales et transnationales d'évaluation contiennent souvent la production ou la diffusion de normes sur l'évaluation, mais elles n'ont pas les moyens de contrôler leur respect par les gouvernements. Seuls les réseaux intégrés dans des organisations internationales, comme Evalnet au CAD-OCDE, ont les moyens de contrôler, *via* des examens par les pairs, le respect par les États membres des « bonnes pratiques » d'évaluation auxquelles ils ont adhérées. Pour les pays du Sud, la seule contrainte provient, non pas des réseaux d'évaluateurs, mais des organisations internationales qui ont fait des « bonnes pratiques » d'évaluation une condition de leur allocation de fonds d'aide au développement[953].

Selon le politiste Richard Common, c'est par la construction d'une élite transnationale que les réseaux travaillant sur des thématiques liées à la bonne gouvernance gagnent une influence normative et politique : « un mouvement d'homogénéisation culturelle et idéologique des élites est actuellement en cours, qui favorise, dans certains secteurs tels que la gestion des politiques publiques, la formation de larges consensus internationaux concernant la définition des problèmes politiques et la détermination de critères de bonne politique ou de bonne pratique. Ce phénomène est accentué par la montée d'une élite transnationale de la coopération pour le développement (constituée de responsables d'ONG transnationales, de cadres d'organismes d'aide, de chercheurs et de consultants) vouée à la diffusion de recettes pour

[952] *Cf.* Chapitre 3.
[953] *Cf.* Chapitre 8.

traiter les enjeux politiques en vogue »[954]. Dans le cas des sociétés d'évaluateurs, si nous voyons bien une élite se constituer dans les pays du Nord et les organismes multilatéraux, cette élite est moins visible au Sud. Il est vrai que certains évaluateurs parviennent à intégrer des réseaux internationaux, à faire reconnaître leur expertise au niveau international et à collaborer pleinement avec des évaluateurs du Nord, par exemple, l'Égyptienne Nermine Wally, la Béninoise Justine Odjoube ou le Sénégalais Moctar Saw, mais ces personnalités restent des exceptions[955].

Enfin, une partie de la littérature sur les réseaux souligne l'importance de la bureaucratisation et de l'institutionnalisation des réseaux pour qu'ils gagnent en influence politique. Peter Haas développe particulièrement cet aspect : « les communautés épistémiques peuvent diffuser leur expertise et influencer les gouvernements nationaux et les organisations internationales en occupant des niches dans les services de régulation et de conseil. Cela suggère que l'influence auprès des décideurs politiques dépend de la capacité des groupes transmettant un savoir à gagner et à exercer un pouvoir bureaucratique »[956]. Selon lui, pour exercer une influence politique, les membres des réseaux nationaux ou transnationaux doivent infiltrer les instances de prise de décision pour y institutionnaliser leur expertise. À ce jour, les cas où les réseaux d'évaluation sont directement et régulièrement intégrés à la prise de décision des gouvernements concernant les réformes de politiques publiques sont encore rares. Les réseaux d'évaluation sont toujours à la recherche de points d'ancrage auprès du pouvoir politique, de « champions de l'évaluation », pour reprendre le vocabulaire de la Banque mondiale, pour faire valoir et relayer leur expertise et espérer que celle-ci soit un jour intégrée à la prise de décision. La plupart des réseaux du Sud restent informels, peu structurés et peu organisés, ce qui peut nuire, comme nous l'avons vu, à leur influence politique.

Notre analyse montre que l'influence des réseaux nationaux et transnationaux d'évaluation est multiple. Premièrement, ils ont une influence cognitive sur leurs membres et sur les gouvernements qui vont développer un discours et une compréhension commune de l'évaluation. C'est le type d'influence le plus facile à acquérir. Deuxièmement, les réseaux ont une influence normative, grâce à laquelle ils vont produire et/ou diffuser des « bonnes pratiques » d'évaluation harmonisées auprès de leurs membres. Ils en font ensuite le plaidoyer auprès des instances gouvernementales avec lesquelles ils ont pu établir des liens. Les réseaux ont également une influence symbolique, dans la mesure où ils organisent régulièrement des

[954] COMMON, Richard, *Public Management and Policy Transfer in Southeast Asia*, Aldershot, 2001, 300p, p. 21.
[955] Leurs parcours académiques et professionnels sont présentés en annexe 10.
[956] HAAS, Peter, 1997, *op. cit.* p. 29.

évènements sur l'évaluation et où ils produisent des publications sur ce thème. Ils permettent ainsi de rendre visible l'enjeu du développement de la pratique de l'évaluation au Sud et d'une approche plus inclusive et égalitaire (envers les acteurs du Sud) des évaluations des programmes de développement. Enfin, les réseaux ont une influence politique. Une fois qu'ils ont créé des relations de longue durée et de confiance avec les décideurs politiques et que leur expertise est reconnue comme utile pour éclairer la prise de décision, les membres du réseau les plus influents vont pouvoir peser sur la gestion des politiques publiques.

Nous pouvons conclure que l'influence normative et politique des réseaux d'évaluation est (encore) faible sur les États ciblés par leurs actions de plaidoyer. Toutefois, ce cas de figure est commun à la majorité des réseaux transnationaux et nationaux. Keck et Sikkink expliquent, par exemple, que « les réseaux contribuent à changer les perceptions que les États et les acteurs sociétaux ont de leurs identités, intérêts et préférences, à transformer leurs positions discursives et finalement à changer les procédures, politiques et comportements »[957]. Nous retrouvons ici l'idée d'une influence qui se réalise sue le long terme, par un changement cognitif, normatif et politique graduel des États et de la société civile.

- **Conclusion de chapitre**

Ces dernières années, les acteurs se mobilisant pour diffuser une culture de l'évaluation dans les pays du Sud se sont dotés d'un nouveau vecteur d'influence : les réseaux d'évaluateurs. Après avoir mené un plaidoyer de plusieurs années auprès de « champions » de l'évaluation dans la plupart des pays du Sud, les évaluateurs des organisations multilatérales, notamment de l'UNICEF, ont soutenu la création de plusieurs dizaines de réseaux d'évaluateurs nationaux et régionaux au Sud. Ces évaluateurs ont également créé des réseaux transnationaux d'évaluateurs financés par des coalitions de donateurs, à la tête desquels les Nations unies. Ainsi, en une décennie, de réseaux d'évaluateurs ont proliférés dans la plupart des pays du monde. Nous avons démontré que ces réseaux, présentés comme un vecteur égalitaire de renforcement des capacités en évaluation et de diffusion d'une culture de l'évaluation dans les pays du Sud, sont en réalité, pour la plupart, sous la tutelle et l'influence des donateurs. Les échanges horizontaux d'évaluateurs du Sud à évaluateurs du Nord sont rares. Ces deux types d'acteurs se méconnaissent et travaillent peu ensemble dans le cadre de l'évaluation des programmes d'aide au développement, alors que des collaborations seraient possibles. Ce sont toujours les mêmes

[957] KECK, Margaret, SIKKINK, Kathryn, 1998, *op. cit.*

évaluateurs du Sud qui se retrouvent associés aux travaux dirigés par les évaluateurs du Nord. Ce constat conduit à remettre en cause l'existence d'un réseau transnational d'évaluateurs qui intégrerait les évaluateurs du Nord, du Sud et les sociétés d'évaluation nationales, régionales et internationales. Si l'existence d'un tel réseau transnational est possible, le processus n'est pas achevé. Les réseaux du Sud étant souvent faibles et leurs contextes d'action relativement défavorables au développement d'une culture de l'évaluation, leur influence normative et politique reste limitée. S'ils permettent un renforcement des capacités des évaluateurs déjà en fonction, ils ont des difficultés à attirer de nouveaux acteurs et à faire reconnaître leur expertise par des gouvernements qui les tolèrent, apprécient la vitrine de bonne gouvernance qu'ils offrent, mais reconnaissent peu l'utilité de leur expertise. Sans un changement politique global dans les pays du Sud, l'influence des réseaux transnationaux restera limitée.

PARTIE III : Les effets de l'évaluation : reclassement, contrôle ou démocratisation ?

Cette troisième et dernière partie a pour objectif d'analyser les effets de la diffusion d'une culture de l'évaluation sur le système d'acteurs de l'aide au développement. Deux aspects seront particulièrement explorés, d'une part les effets de reclassement[958] qu'entraîne l'évaluation sur les acteurs de l'aide, qu'ils soient donateurs ou bénéficiaires et, d'autre part les effets transformateurs qu'entraîne l'évaluation sur les politiques d'aide au développement. Par reclassement, nous entendons les effets de d'inclusion ou au contraire de stigmatisation que connaissent les acteurs au sein du système de l'aide, en fonction de leur conformité aux « bonnes pratiques » internationales de l'évaluation.

Ces thématiques de recherche sont nées du constat que la littérature sur l'aide au développement n'envisage que très rarement la question de l'évaluation sous son angle « relationnel »[959]. Il y a peu d'analyses sur les enjeux de pouvoir et de reclassement qui émergent à partir de l'évaluation. Selon les discours des représentants de l'OCDE[960], l'évaluation est censée rendre plus égales les relations entre donateurs et bénéficiaires en permettant des échanges et une meilleure transparence sur l'efficacité des programmes d'aide. L'évaluation devrait favoriser l'autonomisation des bénéficiaires en leur permettant de s'approprier les programmes d'aide. Si nous exprimons des doutes quant à la capacité actuelle des évaluations de permettre une telle « égalitarisation » des relations entre donateurs et bénéficiaires, il nous semble important de souligner que l'évaluation est une sorte de « jeu de rôle », au sein duquel les stratégies, intérêts et idéologies de chaque partie prenante s'expriment. Or, dans la littérature actuelle, l'évaluation est presque exclusivement abordée sous son aspect technique de mesure de l'efficacité de l'aide et à partir de sa fonction de légitimation et de redevabilité. Pour sortir de cette approche restrictive, nous allons aborder les effets de l'évaluation en nous concentrant sur leurs aspects relationnels et politiques.

[958] Pour une analyse plus large du phénomène de reclassement induit par la mise en place de bonnes pratiques par les organisations internationales, voir KLEIN, Asmara, LAPORTE, Camille, SAIGET, Marie, 2015 (à paraître), *op cit.*

[959] Rosalind Eyben est une des rares chercheurs à développer une étude de la question de l'efficacité de l'aide en termes « relationnels ». EYBEN, Rosalind, "Hiding relations: the irony of 'effective aid "European *Journal of Development Research*, vol. 22, n°3, pp. 382-397. Nous appliquons en partie son approche à notre objet d'étude qu'est l'évaluation de l'aide au développement.

[960] Entretien avec une Senior Policy Advisor du département de l'efficacité de l'aide au CAD-OCDE, 2 juillet 2014, Entretien avec un Senior Policy Advisor du département de l'efficacité de l'aide au CAD-OCDE, 2 juillet 2014.

L'élaboration de l'agenda sur l'efficacité de l'aide et ses exigences en termes d'évaluation participative, de transparence et de redevabilité mutuelle[961] a été en partie motivée par la croyance que leur application allait entraîner une « égalitarisation » des relations entre donateurs et bénéficiaires. Par l'introduction de ces nouvelles techniques dans les pratiques des donateurs et des bénéficiaires, l'objectif du CAD-OCDE était de transformer les relations de pouvoir dans le champ de l'aide au développement. Au problème politique du manque d'autonomie des bénéficiaires de l'aide et de l'inégalité des rapports de force entre donateurs et bénéficiaires, les donateurs ont donc apporté une solution technique, la mise en place de procédures participatives ou mutuelles de suivi-évaluation des programmes d'aide.

Nous voyons ici une volonté des donateurs de dépolitiser l'aide au développement et l'évaluation pour effacer les enjeux de pouvoir qui la structurent. Toutefois, vouloir augmenter l'autonomie des bénéficiaires de l'aide par des techniques dépolitisées nous apparaît peu crédible, cet enjeu étant éminemment politique. Le CAD-OCDE diffuse l'idée selon laquelle, dans la mesure où les évaluations sont participatives et la redevabilité mutuelle, les donateurs et les bénéficiaires de l'aide seraient mis sur un pied d'égalité, chacun devant rendre compte à l'autre de ses efforts et de ses résultats.

De plus, il est faux de voir l'évaluation comme une technique dépolitisée. L'évaluation est une activité porteuse de valeurs politiques (transparence, démocratie participative, etc.) et qui voit se confronter les rapports de force des parties prenantes des programmes d'aide. L'évaluation n'a de sens que dans l'usage qui en sera fait par les acteurs des politiques publiques et formule des recommandations qui servent à réorienter les programmes et politiques évalués. La diffusion d'une image dépolitisée de l'évaluation a pour objectif de faciliter le transfert des « bonnes pratique » de l'évaluation, et plus largement de l'efficacité de l'aide, aux nouveaux acteurs du développement, tels que les donateurs émergents, les fondations, le secteur privé, etc. En effet, lorsque les aspects politiques de l'évaluation sont effacés derrière la technique, ces « bonnes pratiques » apparaissent plus facilement transposables dans des contextes politiques et organisationnels divers.

Notre hypothèse de départ était que la multiplication des « bonnes pratiques » d'évaluation n'a pas eu l'effet escompté d'amélioration de l'égalité dans les relations entre donateurs et bénéficiaires de l'aide. De même, nous pensons que l'évaluation n'a pas permis de rationaliser les politiques d'aide, mais est plutôt l'instrument de leur légitimation. Nous

[961] Voir lexique en annexe 7.

estimons que l'évaluation, telle qu'elle est mise en œuvre aujourd'hui, conduit à technocratiser la mise en œuvre de l'aide au développement et à creuser le différentiel de compétences et d'expertise en faveur des donateurs qui sont les seuls à disposer (ou à se prévaloir de disposer) des compétences de l'évaluation (statistiques, économétrie, enquêtes en sciences sociales, etc.). D'autre part, nous faisions l'hypothèse en introduction que la montée en puissance d'une pratique normalisée de l'évaluation entraînerait des effets de reclassement et de résistance aussi bien parmi les donateurs que parmi les pays bénéficiaires de l'aide au développement.

Nous explorerons ces hypothèses dans trois chapitres en nous attachant à décrypter tous les enjeux politiques qui se dissimulent derrière l'utilisation de l'évaluation dans le secteur de l'aide au développement.

Dans un premier chapitre, nous analyserons les effets qu'entraîne la suprématie actuelle des « bonnes pratiques » de l'efficacité de l'aide, et plus particulièrement de l'évaluation, sur les politiques d'aide des donateurs émergents. En effet, ces donateurs ont une approche de l'aide au développement où la culture de l'évaluation est peu (voire pas) développée et où l'efficacité de l'aide répond à des critères qu'ils présentent comme étant très différents de ceux des donateurs traditionnels. Nous nous interrogerons sur la croyance répandue chez les donateurs occidentaux selon laquelle les pratiques des donateurs émergents menaceraient leur modèle d'efficacité de l'aide. Nous analyserons le jeu d'acteurs qui se construit entre donateurs traditionnels et donateurs émergents autour de la question de l'intégration de ces derniers dans le système normatif actuel de l'efficacité de l'aide. En effet, en instituant un cadre de bonnes pratiques, les donateurs occidentaux ont créé une dichotomie entre ceux qui les respectent et ceux qui ne les respectent pas. En conséquence, les seconds se retrouvent « déclassés » par rapport aux premiers, parce qu'ils sont considérés comme de « mauvais donateurs ». Mais quels sont les effets de ce déclassement des donateurs émergents sur la scène internationale ? Nous questionnerons également l'existence d'effets d'apprentissage, ou d'émulation, entre ces deux catégories de donateurs.

À cet effet, nous prendrons comme cas d'étude, la Chine et la Corée du Sud. La Chine car il s'agit du donateur émergent le plus puissant en termes de volumes d'aide et le plus visible dans les dialogues avec les donateurs traditionnels et la Corée du Sud car elle a adhéré en 2011 au CAD-OCDE adoptant alors, contrairement à la Chine, une stratégie d'intégration des bonnes pratiques des donateurs traditionnels dans sa politique d'aide au développement. Nous analyserons les logiques qui sous-tendent les choix de ces deux donateurs et les transformations

qui s'opèrent en conséquence dans leurs politiques et pratiques d'aide, spécifiquement en matière d'évaluation.

Enfin, nous verrons dans quelle mesure, la présence croissante, sur le terrain de l'aide, de donateurs émergents aux exigences plus souples en termes de suivi et d'évaluation concurrence les donateurs traditionnels. Nous essaierons d'apprécier la réalité de l'idée selon laquelle les gouvernements des pays bénéficiaires de l'aide se détourneraient de l'aide des bailleurs du CAD-OCDE en faveur de l'aide des donateurs émergents, réputée moins ingèrente et plus équitable. Si un tel déclassement des donateurs occidentaux s'avère réel, il y aurait alors un rapport de force inversé au niveau international et au niveau local entre donateurs traditionnels et émergents.

Dans un deuxième chapitre, nous étudierons les effets de l'évaluation sur l'allocation de l'aide au développement. Notre questionnement est de savoir dans quelle mesure l'évaluation de l'efficacité de l'aide influence la répartition des fonds d'aide par les donateurs. L'enjeu sous-jacent est celui d'un éventuel reclassement des pays bénéficiaires dans les listes de pays prioritairement bénéficiaires de l'aide, en fonction de leur adéquation aux critères d'évaluation retenus par les donateurs. Nous démontrerons que les bénéficiaires de l'aide sont davantage « objet » que « sujet » des évaluations et que, dans un contexte où aide au développement et sécurité internationale sont de plus en plus liées, l'évaluation devient un outil de « gouvernementalité » des populations du Sud. Nous développerons ainsi une analyse foucaldienne du rôle de l'évaluation dans l'aide au développement, en questionnant son rôle d'instrument de contrôle des populations bénéficiaires et de leurs gouvernements[962]. Nous nourrirons notre argumentation de notre expérience de terrain au Laos lors d'une observation participante au sein de l'agence locale de l'AFD. Pour compléter cette analyse en termes de « gouvernementalité », nous nous demanderons ensuite comment les bénéficiaires de l'aide pourraient se saisir du suivi-évaluation pour mieux contrôler les projets d'aide et faire entendre leur voix dans le système de l'aide. La dernière partie de ce chapitre sera consacrée à la remise en cause de l'approche limitative de l'évaluation comme un instrument de domination et à l'analyse des tentatives de réappropriation de l'évaluation par les bénéficiaires de l'aide.

Dans un dernier chapitre, nous explorerons les effets de reclassement que produisent des évaluations de la qualité des politiques d'aide des donateurs. Durant de nombreuses décennies,

[962] Pour une analyse plus générale des phénomènes de dépolitisation et de gouvernementalité dans le secteur de l'aide au développement voir : FERGUSON, James, *The Anti-Politics Machine: Development, Depoliticization and Bureaucratic Power in Lesotho,* University of Minnesota Press, 1994, 336p.

les politiques d'aide des donateurs étaient peu évaluées, et ce malgré la montée en puissance de l'évaluation dans le secteur de l'aide au développement. Alors que les pays bénéficiaires de l'aide se voyaient contraints d'évaluer leurs politiques publiques nationales pour prouver aux donateurs le bon usage qu'ils faisaient de l'aide au développement qui leur était fournie, les donateurs n'avaient pas à rendre compte de la qualité de leurs politiques et pratiques d'aide.

À la suite d'un plaidoyer de la société civile et en raison de l'exigence de redevabilité de plus en plus forte pesant sur les donateurs, ces derniers ont commencé à mettre en place des mécanismes d'évaluation par les pairs de leurs politiques d'aide au développement, particulièrement de celles des organisations multilatérales. De plus, la société civile s'est mise à réaliser des évaluations de même type. Abordant l'évaluation comme un objet politique, nous démontrerons que ces évaluations des politiques d'aide des donateurs sont porteuses d'une approche particulière de l'aide au développement, celle de l'agenda sur l'efficacité de l'aide puis du Partenariat global pour le développement. Nous démontrerons que ces évaluations conduisent à déclasser et stigmatiser les donateurs qui ne se conforment pas à cette approche de l'aide et qu'elles ont pour objet de les inciter à modifier leurs politiques d'aide. Nous analyserons si cet effet incitateur est efficace et dans quelle mesure il mène à une harmonisation des mandats des organismes d'aide au développement.

Chapitre 7 : Un régime de l'aide fragmenté autour des « bonnes pratiques » d'évaluation : le catalyseur des donateurs émergents

Un des changements ayant affecté le régime de l'aide au développement durant la dernière décennie est la montée en puissance de donateurs non-membres du CAD-OCDE, notamment les fonds verticaux, ONG, acteurs privés et donateurs émergents. Dans ce chapitre, nous nous focalisons sur les donateurs émergents, en raison des positionnements et des stratégies spécifiques qu'ils tiennent face aux « bonnes pratiques » d'évaluation et plus largement d'efficacité, prônées par les donateurs dits traditionnels.

A. Des effets de déclassement sur la scène internationale

Le terme « donateurs émergents » [963] regroupe des États dont le volume d'aide fourni augmente fortement et/ou des pays qui sont à la fois bénéficiaires et donateurs. Ce terme est toutefois trompeur, car certains des pays désignés comme tels, notamment la Chine, ont une activité d'aide au développement depuis de plusieurs décennies. La Chine a commencé à mettre en place des projets d'aide dans les années 1950 qui se concentraient alors en Asie (notamment au Viêtnam et en Corée du Nord)[964]. C'est le redéploiement, géographique et en termes de volume, de l'aide chinoise dans les années 2000, qui fait qualifier cet État de « donateur émergent ». Certains préfèrent parler de donateurs « non-CAD-OCDE » ou « du Sud ». Nous avons choisi le terme « donateur émergent », car nous incluons la Corée du Sud dans cette catégorie, pays qui fait partie du CAD-OCDE depuis 2011 et qui n'est pas un pays « du Sud ».

Il existe peu d'estimations globales des volumes d'aide fournis par les donateurs émergents, étant donné que certains, notamment la Chine et l'Inde, ne diffusent pas d'informations quantitatives et qualitatives sur leurs actions d'aide au développement auprès du CAD-OCDE. Pour pallier ce manque, nous avons fait des recherches complémentaires, afin de trouver des estimations crédibles des montants d'aide alloués par ces deux pays, que nous avons additionnés à ceux recensés par le CAD-OCDE pour les autres donateurs émergents. Ainsi, nous estimons que le montant d'aide fourni par les donateurs émergents atteignait 12

[963] Les donateurs émergents regroupent des membres de l'Europe (Chypre, République Tchèque, Estonie, Hongrie, Lettonie, Lituanie, Pologne, Roumanie, Slovaquie, Slovénie ; d'autres donateurs « du Nord » (Corée du Sud, Islande, Israël, Lichtenstein, Turquie, Russie) des fournisseurs de coopération Sud-Sud (Chine, Inde, Brésil, Taïwan, Thaïlande, Koweït, Arabie Saoudite, Émirats Arabes Unies).
[964] China's Information Office of the State Council, *White Paper on China's Foreign Aid,* 21 Avril 2011.

milliards de dollars en 2010[965]. Ce volume d'aide représentait environ un dixième de l'aide fournie la même année par les pays membres du CAD-OCDE[966]. L'aide de la Chine constituait environ 20 % de ce total.

- **La stratégie inclusive du CAD-OCDE envers les donateurs émergents**

Jusqu'à aujourd'hui, les acteurs prédominants du régime international de l'aide ont été les donateurs bilatéraux membres du CAD-OCDE et les organisations multilatérales. Ces acteurs ont défini les actuelles « bonnes pratiques » de l'efficacité de l'aide au développement. Elles sont directement influencées par l'idéologie du New Public Management qui fait de l'évaluation un outil central de gestion des politiques publiques. Les donateurs occidentaux ont conscience que ces « bonnes pratiques » sont l'émanation d'une vision occidentale de la bonne gouvernance qui n'est pas partagée par les donateurs émergents. C'est pourquoi ils craignent que ces derniers sapent leurs efforts d'instituer ces « bonnes pratiques » comme le modèle à suivre par tous les acteurs du régime de l'aide au développement[967]. Nous observons qu'une « fracture identitaire » se créée parmi les donateurs, entre ceux qui adhèrent aux principes de l'efficacité de l'aide et ceux qui n'y adhèrent pas[968]. Selon nous, cette fracture est plus visible et plus profonde en ce qui concerne les « bonnes pratiques » d'évaluation, de transparence et de redevabilité qui divisent davantage les donateurs « émergents » et « traditionnels », qu'en ce qui concerne les « bonnes pratiques » d'appropriation ou d'alignement, qui font plus facilement l'unanimité en théorie, bien qu'elles soient peu appliquées en pratique.

Ce type de polarisation autour d'un ensemble de bonnes pratiques internationales est décrit par David Lumsdaine : « adopter et incarner les normes d'un régime en utilisant les « mots clés » dans ses politiques et ses directives peut, bien sûr, n'être que pure rhétorique. Toutefois, même dans ce cas, cela montre leur désir d'« être vus » comme incarnant les principes et les normes qui constituent le régime international de l'aide au développement »[969]. Cette citation

[965] HUANG, Meibo, REN, Peiqiang, « L'aide étrangère de la Chine dans l'architecture de l'aide internationale », *Revue internationale de politique de développement*, vol. 3, 2012, pp. 1-18 ; SMITH, Kimberley, YAMASHIRO, Tallta, ZIMMERMANN, Felix, *Beyond DAC, the Welcome role of other providers of development cooperation*, OECD Development Cooperation Directorate, mai 2012.

[966] http://stats.oecd.org/qwids/ consulté le 9 avril 2014.

[967] MANNING, Richard, "Will Emerging Donors' Change the Face of International Co-operation? " *Development Policy Review*, vol. 24, n°4, juillet 2006, pp. 371-385; DAVIES, Penny, *Aid Effectiveness and non-DAC providers of Development Assistance*, Informal Working Group on non-DAC Providers of Development Assistance, Background document to Round Table 9 of the Third High Level Forum on Aid Effectiveness, août 2008, GRIMM *et al*., *European Development Cooperation to 2020: Challenge by New Actors in International Development*, Working Paper, EDC 2020, 2009,

[968] Entretien avec une ancienne Chief Social Development Advisor au DFID, consultante auprès d'organismes de développement, 14 février 2014.

[969] LUMSDAINE, David. H., *Moral Vision in International Politics: The Foreign Aid Regime 1949-1989*, Princeton University Press, 1993, 376p, pp. 172-228.

exprime bien la portée normative de l'instauration de bonnes pratiques apparemment neutres et techniques, comme celles de l'efficacité de l'aide. En effet, les donateurs émergents sont souvent ceux qui ne respectent pas, ou peu, les bonnes pratiques du CAD-OCDE. En conséquence, les donateurs émergents sont catégorisés comme étant des « mauvais donateurs ». Ils se trouvent donc indirectement déclassés et stigmatisés par l'instauration de la Déclaration de Paris qui formalise la différence entre les « bons » et les « mauvais » donateurs. Ce sentiment de déclassement est révélé par la littérature académique postcoloniale sur l'aide au développement, très représentée dans les pays émergents[970]. Kapoor, chercheur indien, condamne l'exclusion des donateurs émergents du « clan » du CAD. Il estime que le travail normatif du CAD est « sélectif et insulaire, discriminant envers les donateurs du Sud ». Cette analyse, qui pousse à l'extrême le phénomène de déclassement qu'induit la production de bonnes pratiques, a le bénéfice de pointer leur caractère normatif. Par leur qualification même, les « bonnes » pratiques contiennent un jugement de valeur sur toutes les pratiques alternatives, qui sont délégitimées. En conséquence, certains donateurs ont cherché à rejoindre le « club » des bons donateurs pour se reclasser (la Corée du Sud) dans le régime international de l'aide, alors que d'autres ont adopté une stratégie de résistance et de remise en cause des bonnes pratiques portées par les donateurs traditionnels (la Chine).

Les donateurs traditionnels ont, quant à eux, pris conscience qu'ils n'avaient plus le monopole de la production de bonnes pratiques en matière de développement. Homi Kharas observe : « tandis que le CAD-OCDE reste le cœur du système global de l'aide, son monopole de l'APD s'érode avec l'émergence des nouveaux partenaires de développement (…) Les donateurs traditionnels qui forment le CAD-OCDE ne peuvent plus déclarer qu'ils parlent au nom de l'ensemble de la communauté mondiale des donateurs »[971]. Ce constat nous amène à nous demander quelle stratégie le CAD-OCDE a adopté pour maintenir, voire redéployer, son mandat de producteur de normes sur le développement face à la concurrence de certains donateurs émergents qui offrent d'autres approches de l'aide.

Nous observons que les donateurs traditionnels ont choisi d'adopter une stratégie inclusive par rapport aux donateurs émergents, alors qu'ils les avaient ignorés durant les décennies précédentes. Ce changement brutal de positionnement a largement été commenté dans la littérature sur l'aide au développement. Par exemple, Felix Zimmermann et Kimberley Smith expliquent que la communauté de l'aide *mainstream* est très rapidement passée d'une

[970] Voir par exemple: MCEWAN, Cheryl, *Postcolonialism and development*, Routledge, 2009, 376p.
[971] KHARAS, Homi, *Catalysing Development: A new Vision for Aid*, Brookings, http://www.brookings.edu/events/2011/06/28-development-aid; (retransmission audio en ligne) consulté le 13 mars 2014.

attitude de négligence et de mépris envers les donateurs émergents à des efforts intenses, à partir de 2005, pour établir des relations plus équitables et respectueuses avec eux[972]. Ainsi, en 2005, le Brésil, l'Inde et la Chine, ainsi que d'autres pays à revenus intermédiaires, ont été invités, à titre d'observateurs, au Haut Forum de Paris sur l'efficacité de l'aide. Symboliquement, le CAD-OCDE montra alors qu'il reconnaissait que les États émergents faisaient partie de la « communauté » des donateurs. Toutefois, cette demi-mesure vers l'intégration des donateurs émergents aux négociations de l'agenda sur l'efficacité de l'aide fut mal perçue par ces derniers. Selon Rosalind Eyben, « ils perçurent l'agenda comme étant tourné exclusivement vers les relations traditionnelles de l'aide. Le Brésil déclara ne pas avoir signé la Déclaration de Paris en raison de son caractère Nord-Sud »[973].

En janvier 2009, le Groupe de travail sur l'efficacité de l'aide du CAD-OCDE a créé un Groupe d'étude sur la Chine qui réunit des représentants du CAD et de la Chine pour échanger sur les expériences de chacun en matière de gestion de l'aide au développement et entraîner un apprentissage mutuel. Cette initiative provient notamment des membres britanniques et allemands du CAD qui sont particulièrement engagés dans le renforcement des échanges avec le donateur chinois[974]. Par exemple, le DFID a demandé à ses agences de terrain en Afrique de « construire des liens » avec leurs contreparties chinoises. Il a aussi invité le ministre chinois chargé de la coopération économique et de l'aide au développement à envoyer un observateur pour participer à un examen par les pairs d'un programme d'aide du DFID organisé par le CAD-OCDE[975]. Toutefois, le DFID reste réticent quant à la mise en œuvre de programmes de « coopération triangulaire »[976] avec les donateurs émergents, alors que l'Allemagne, les donateurs nordiques et le Japon y sont plus favorables[977].

Nous pouvons nous demander quel est l'intérêt pour le CAD-OCDE d'instaurer un dialogue avec la Chine, alors que ce pays rejette ouvertement les bonnes pratiques promues par le Comité et affiche sa réticence à s'engager dans des accords liants avec les donateurs traditionnels. Nous avons identifié plusieurs raisons : une volonté de s'informer sur l'aide chinoise qui reste peu transparente ; une volonté d'échanger sur le développement chinois lui-même et sur les leçons qui pourraient être utiles à d'autres pays du Sud ; et surtout, le désir

[972] SMITH, Kimberley, ZIMMERMANN, Felix, "More Actors, More Money, More Ideas for International Development Cooperation", *Journal of International Development,* vol. 23, n°5, juillet 2011, pp. 722-738.
[973] EYBEN, Rosalind, 2012, *op. cit.* p. 459.
[974] Entretiens au siège de l'OCDE.
[975] Entretien avec un membre du DFID, 7 janvier 2013.
[976] L'aide « triangulaire » est une forme d'aide où un donateur traditionnel s'allie à un donateur émergent dans le cadre d'un programme d'aide dans un pays du Sud. On parle aussi d'aide tripartite.
[977] Constat fait lors d'un séminaire interne à l'AFD sur la coopération entre donateurs traditionnels et donateurs émergents qui réunissait des représentants de l'AFD, du DFID et de la JICA, le 28 novembre 2014.

d'exposer les fonctionnaires et les chercheurs chinois aux principes occidentaux de l'aide au développement. Nous notons, en effet, chez les membres du Groupe d'études sur la Chine l'idée selon laquelle exposer les Chinois aux « bonnes pratiques » internationales de suivi-évaluation les incitera à les adopter[978].

En 2009, une *Task Team* sur la Coopération Sud-Sud a également été créée au CAD-OCDE, afin de comprendre les dynamiques de ce type de coopération, d'en identifier les bonnes pratiques et d'explorer sa complémentarité avec la coopération Nord-Sud[979]. Ensuite, la Corée du Sud a adhéré au CAD-OCDE en 2010. Elle a ainsi été le premier donateur émergent à rejoindre formellement le club des donateurs traditionnels.

Richard Manning, ancien directeur du CAD de 2003 à 2007, et en quelque sorte le père de la Déclaration de Paris, témoigne de cette ouverture du CAD, à laquelle il a personnellement travaillé : « j'ai encouragé le renforcement des liens avec le Groupe de coordination des donateurs arabes, qui nous invita à une rencontre au Koweït en février 2007. Le CAD planifie de reconduire ce type de rencontres sur une base régulière et plus formelle. Je me suis personnellement rendu en Chine, en Inde et en Thaïlande pour établir des contacts. Les donateurs émergents ont participé à plusieurs évènements organisés par le CAD. Nous n'avons plus qu'à voir l'ampleur de leur intérêt réel pour ce dialogue »[980]. Cette dernière phrase laisse entrevoir le scepticisme des donateurs traditionnels quant à la volonté réelle des donateurs émergents, notamment la Chine, de s'engager dans un processus d'apprentissage mutuel et de coordination de l'aide avec les membres du CAD-OCDE.

Selon certains chercheurs, l'ouverture des donateurs traditionnels envers les donateurs émergents cache, en réalité, une volonté de sécuriser leur hégémonie sur le régime international de l'aide, tout en paraissant ouverts et inclusifs. Cette ouverture serait vue comme un moyen de pousser les donateurs émergents à se conformer aux bonnes pratiques des donateurs traditionnels et non pas d'établir un réel apprentissage mutuel[981]. Ce serait davantage la méfiance que la confiance qui motiverait les rapprochements observés. Nous rejoignons en partie ces affirmations. Selon nous, l'ouverture du CAD-OCDE aux pays émergents et la création de nouvelles plateformes pour engager un dialogue avec eux sont effectivement des

[978] Entretien avec un analyste des politiques au sein du département de l'efficacité de l'aide du CAD-OCDE, chargé de la coopération avec les donateurs du Sud, 2 juillet 2014.
[979] http://www.oecd.org/dac/effectiveness/taskteamonsouth-southco-operation.htm, consulté le 30 mars 2014.
[980] MANNING, Richard, *The DAC as a Central Actor in Development Policy Issues: Experience over the Past Four Years*, Discussion Paper, Deutsches Intitut für Entwicklungspolitik, German Development Institute, juillet 2008.
[981] SCHULZ, Nils-Sjard, *From Accra to 2011: perspective for the global governance of aid*, FRIDE Comment, Madrid, Fundacion para las relaciones internationals y el Dialogo Exterior, 2008*;* MAWDSLEY, Emma, *From Recipients to Donors: Emerging donors and the Changing Development Landscape*, Zed Books, 2012, 288p.

moyens pour les bailleurs traditionnels de sensibiliser les donateurs émergents aux « bonnes pratiques » actuelles de l'efficacité de l'aide et de les leur faire adopter. Toutefois, les donateurs émergents ont su saisir cette opportunité de faire entendre leur voix et en tirent une plus grande influence sur la production de ces « bonnes pratiques ». Les émergents ne sont donc pas purement dominés ou instrumentalisés, ils tirent un intérêt de leur dialogue avec les donateurs du CAD-OCDE. Cette influence des donateurs émergents a notamment été visible lors du Haut Forum de Busan, évènement que nous allons désormais étudier.

- **Le Forum de Busan : la fin du paradigme de l'efficacité de l'aide au développement ?**

Notre suivi du Forum de Busan et les entretiens que nous avons pu mener auprès de donateurs émergents et de membres du CAD-OCDE, nous ont fait reconsidérer à la hausse le poids des donateurs émergents sur la réforme actuelle des « bonnes pratiques » de l'efficacité de l'aide, par rapport à nos hypothèses de départ et à la littérature existante. Lors de ce forum, les donateurs émergents ont fait évoluer l'agenda politique du régime de l'aide au développement, notamment en reléguant au second plan les questions de la diffusion d'une culture de l'évaluation et du résultat.

Nous avons suivi les négociations et les avancées du Forum à distance via la retransmission sur Internet de certaines sessions de travail et grâce aux réseaux sociaux (Twitter, blogs), ce qui nous a permis d'avoir une bonne idée de l'atmosphère qui y régnait, des coalitions d'acteurs qui se créaient et des enjeux qui émergeaient. De plus, nous avons pu rentrer en contact avec plusieurs participants au Forum qui ont pu compléter notre analyse pat leur expérience.

Le Forum de Busan s'est tenu du 29 novembre au 1er décembre 2011 et a rassemblé plus de 3 000 personnes (donateurs traditionnels, donateurs émergents, OSC, secteur privé, journalistes, etc.). Le Forum devait être consacré au suivi des progrès vers l'atteinte des objectifs de la Déclaration de Paris, ainsi qu'à la réflexion sur l'avenir de cette déclaration. Or, dès le premier jour du Forum, il était clair que l'enjeu prévalent serait en réalité de réussir à trouver une nouvelle approche commune de l'aide au développement intégrant les pays émergents, et surtout la Chine, avec qui un accord semblait particulièrement difficile à établir. Pour les donateurs du CAD-OCDE, l'inclusion de la Chine dans l'accord final semblait être

l'objectif primordial du Forum de Busan, alors que sur les réseaux sociaux, les participants au Forum exprimaient leur scepticisme quant à la possibilité d'atteindre cet objectif[982].

Nous estimons que la Chine et l'Inde ont joué de l'importance donnée à leur ratification de l'accord final pour affirmer leur positionnement alternatif dans le régime de l'aide et afficher une prise de distance avec les donateurs traditionnels. Le comportement de la délégation chinoise est révélateur. En effet, elle n'a participé qu'aux sessions du troisième jour traitant de la coopération Sud-Sud, restant absente des discussions formelles le reste du temps[983]. Les délégations indienne et russe se sont également peu impliquées dans les négociations.

Toutefois, les entretiens que nous avons menés auprès de participants au forum nous ont permis de nuancer cette impression. En effet, les donateurs émergents ont été actifs « derrière la scène », dans les lieux informels de discussion : couloirs, cafétéria, temps de pause, etc. Un ministre africain a eu une réflexion éclairante à ce sujet le deuxième jour du forum : « la Chine et l'Inde aiment jouer « derrière la scène » et restent discrets. Ils sont invisibles, mais ils sont là. Pour finaliser la déclaration finale du forum, des personnes les contactent »[984]. En outre, cette stratégie n'a pas été celle de tous les donateurs émergents. Par exemple, le Brésil, le Mexique, ou encore la Corée du Sud, ont participé activement et visiblement aux sessions de travail pour affirmer leur rôle de chefs de file dans le développement d'une coopération Sud-Sud pour le développement. Ainsi, Marco Farani, directeur de l'agence brésilienne de développement, a activement mis en avant les bénéfices de la coopération Sud-Sud et le rôle du Brésil dans cette coopération : « la coopération Sud-Sud du Brésil est fondée sur le partage de bonnes pratiques, de technologies et de compétences dans plusieurs domaines où le Brésil a progressé, spécifiquement l'agriculture et le développement social. Le Brésil soutient une réelle coopération Sud-Sud, durable, détachée de toutes conditionnalités, déterminée par la demande, et qui donne une grande place à l'échange de bonnes pratiques »[985]. Les représentants de la Corée du Sud se sont, quant à eux, positionnés comme un « pont » entre les donateurs

[982] Quelques blogs et réseaux sociaux consultés : Center on Foreign Relations : http://www.cfr.org/; The Broker http://www.thebrokeronline.eu/; Open Democracy : http://www.opendemocracy.net/; DEVEX https://www.devex.com/en/, consultés en décembre 2011.
[983] KIM, Sung-Mi, MAWDSLEY, Emma, SAVAGE, Laura, "A post aid world? Paradigm shift in foreign aid and development cooperation at the 2011 Busan High Level Forum", *The Geographical Journal*, vol. 180, n°1, mars 2014, pp. 27-38.
[984] Entretien avec une ancienne Chief Social Development Advisor au DFID, consultante auprès d'organismes de développement, 14 février 2014.
[985] Marco Farani, allocution au Forum de Busan, section thématique sur la coopération Sud-Sud et triangulaire, 29 novembre 2011.

traditionnels et les donateurs émergents, dont le rôle serait de faciliter le dialogue et la compréhension entre ces deux catégories d'acteurs[986].

Nous estimons donc que l'influence des donateurs émergents dans le Forum a été importante, même si elle s'est souvent exercée en « off » des sections thématiques officielles. Les donateurs émergents ont réussi à évincer les questions de l'évaluation et de la gestion axée sur les résultats de l'agenda du forum, enjeux qui devaient pourtant être en haut des priorités de négociation. L'influence des donateurs émergents, spécifiquement de la Corée du Sud, s'est faite ressentir dès les préparatifs du forum. Selon Rosalind Eyben, la Corée du Sud a pris conscience, après son entrée au CAD, que les membres du Comité traitaient la Déclaration de Paris comme si elle était « périmée », essayant de se désengager des objectifs qu'ils n'avaient pas réussi à atteindre à ce jour, notamment en se servant du fait que les donateurs émergents ne suivaient pas ces principes, qui ne seraient alors « plus d'actualité » ni « réalisables »[987]. En effet, l'évaluation de la Déclaration de Paris a révélé qu'une seule des 13 cibles fixées avait été atteinte en cinq ans et a ainsi dévoilé les faibles progrès des donateurs. Les aspects ayant le moins progressé étant la gestion axée sur les résultats et la redevabilité mutuelle, soit ceux dont l'atteinte était en grande partie conditionnée par une diffusion de la culture de l'évaluation dans les organismes d'aide et les pays bénéficiaires[988].

Ainsi, lors des préparatifs du Forum de Busan, la Corée du Sud a proposé qu'au lieu de se concentrer sur le suivi des résultats de la Déclaration de Paris, comme le faisaient les Hauts Forums précédents, ce Forum s'intéresse à la construction d'un « partenariat global » pour le développement, c'est-à-dire d'un régime de l'aide plus égalitaire. L'argument principal de la Corée du Sud était que le CAD-OCDE ne parviendrait pas à convaincre les autres donateurs émergents de participer au Forum de Busan si l'agenda se concentrait sur l'évaluation de la Déclaration de Paris ; déclaration contre laquelle beaucoup s'étaient élevés[989]. Selon nous, cette requête arrangeait les membres du CAD qui y voyaient un moyen d'éviter de faire publiquement un mauvais bilan de l'agenda sur l'efficacité de l'aide.

Durant la session plénière d'ouverture du Haut Forum de Busan, Brian Atwood, alors directeur du CAD, a rapidement fait référence aux résultats médiocres de l'évaluation de la

[986] LEE, Seong Hoon, HONG Moon Suk, KANG Hanee, *The Outcomes and Future Challenges of the Busan HLF-4: CSO's Perspectives*, Issue Brief 2, KOFID, REDI, février 2012.
[987] EYBEN, Rosalind, 2013, *op. cit*. pp. 459-460.
[988] WOOD *et al, L'évaluation de la Déclaration de Paris, Phase 2 rapport final,* DAC-OECD, mai 2011.
[989]Entretien avec une ancienne Chief Social Development Advisor au DFID, consultante pour des organismes de développement, 14 février 2014.

Déclaration de Paris[990], qui n'ont plus été abordés durant le Forum. Selon Emma Mawdsley, Seung-Mi Kim et Laura Savage, « les questions posées par le président de la section thématique sur la redevabilité à propos des raisons de l'absence de progrès sur cette thématique sont restées sans réponse. L'évaluation de la Déclaration de Paris n'a pas été discutée. Le responsable de cette évaluation, Bernard Woods, n'a même pas été appelé à présenter ses résultats, ni à participer à un panel. (…) Paradoxalement, ce sont les pays bénéficiaires qui ont demandé que les engagements de résultats des Hauts forums de Paris et d'Accra soient reconduits dans la Déclaration finale de Busan. Ils y voyaient un moyen de maintenir une garantie de redevabilité des donateurs sur leurs programmes d'aide »[991]. En effet, un aspect peu évoqué sur l'agenda de l'efficacité de l'aide est que les « bonnes pratiques » d'évaluation participative, de redevabilité mutuelle, d'utilisation des mécanismes de suivi-évaluation des pays bénéficiaires, etc. sont autant de moyens pour les pays du Sud d'inciter les donateurs à fournir une aide de qualité et à les inclure dans la mise en œuvre des programmes. Alors que la Déclaration de Paris est souvent présentée comme un accord technocratique renforçant la domination des donateurs sur les bénéficiaires de l'aide, nous voyons ici que les bénéficiaires la considéraient, dans une certaine mesure, comme un cadre protégeant leurs intérêts.

La principale victoire des donateurs émergents durant le Forum de Busan a été la réussite de leur plaidoyer en faveur d'un abandon de l'expression « efficacité de l'aide », remplacée par « efficacité du développement ». L'enthousiasme des acteurs de l'aide à propos de ce nouveau paradigme a été à la suite du forum de Busan. Sur les réseaux sociaux, les participants, qu'ils soient donateurs, bénéficiaires de l'aide, membres de la société civile, ou représentants du secteur privé, parlaient de « la fin de l'aide au développement », d'une « nouvelle ère du développement », ou encore de la fin de « l'aide postcoloniale »[992]. Cette idée d'un tournant idéologique a été véhiculée par des personnalités politiques éminentes. Par exemple, Tony Blair, ancien premier ministre du Royaume-Uni, déclara que « c'est la première fois que nous pouvons percevoir la fin de l'aide d'ici une génération »[993]. Même le président du CAD-OCE, Brian Atwood, enterra le principe de l'aide : « Dieu merci, même le terme « aide » finira par mordre la poussière »[994]. Après leur participation à Busan, Eng-Mi Kim, Emma Mawdsley et Laura Savage ont noté que « pour la première fois, le mot « aide » est devenu péjoratif dans le discours public sur le développement. Busan était constamment présenté par ses participants

[990] Brian Atwood, allocution lors de la session plénière du Forum de Haut Niveau de Busan, 30 novembre 2011.
[991] KIM, Sung-Mi, MAWDSLEY, Emma, SAVAGE, Laura 2014, *art. cit.* p. 31.
[992] Par exemple, DEVPOLICY : http://devpolicy.org/busan-a-new-global-divide-on-aid20111202/ consulté le 2 décembre 2011 ; Development Gateway : http://www.developmentgateway.org/
[993] KIM, Sung-Mi, MAWDSLEY, Emma, SAVAGE, Laura, 2014, *art. cit.* p. 30.
[994] Brian Atwood, allocution lors de la séance plénière du Haut forum de Busan, 30 novembre 2011.

comme la fin de l'agenda sur l'efficacité de l'aide et comme un nouveau départ vers la poursuite d'une plus large « efficacité du développement »[995]. C'est autour de ce nouveau paradigme de l'efficacité du développement que s'est joué le succès du Haut Forum de Busan. En effet, les donateurs émergents ont fait savoir qu'ils ne ratifieraient la déclaration finale que si elle marquait une sortie de l'aide traditionnelle vers une vision plus égalitaire de la coopération internationale, fondée sur le respect de la souveraineté nationale et le partenariat. Ainsi, nous observons que les discussions autour des bonnes pratiques « techniques » de l'agenda de l'efficacité ont été le catalyseur d'un débat politique sur l'aide au développement.

Nous considérons que le succès du concept « d'efficacité du développement » a été un succès, car il a permis un compromis entre les donateurs traditionnels et émergents en dépassant la notion d'aide, condamnée par les donateurs émergents et certains bénéficiaires comme étant postcoloniale et inégalitaire, tout en conservant l'idée d'efficacité, et donc de résultats.

Ce n'est donc plus l'aide qui doit être efficace et évaluée, mais les politiques de développement dans leur ensemble, rendant encore plus complexe la tâche d'évaluation des activités de coopération pour le développement. La déclaration finale de Busan définit les caractéristiques de l'efficacité du développement : « une croissance forte, durable et partagée », « l'intégration des pays en développement dans l'économie mondiale », « un renforcement de la coopération Sud-Sud », « la participation plus active du secteur privé et la mobilisation de fonds privés », « la facilitation de l'investissement » et « l'appui aux institutions gouvernementales »[996]. Ces caractéristiques reflètent les pratiques actuelles des donateurs émergents qui se concentrent sur la croissance économique et la création de richesse, en investissant dans les secteurs productifs (industrie, infrastructure) et en s'appuyant sur le secteur privé. Ceci contraste avec l'approche des donateurs traditionnels qui mettent l'accent sur le renforcement des secteurs sociaux (santé, éducation) pour réduire la pauvreté et travaillent au renforcement des institutions publiques des pays du Sud (ministères, administrations publiques, gouvernement). Nous comprenons ainsi que la Déclaration de Busan a largement été influencée par les donateurs émergents, comme en témoignent les références multiples à l'importance du renforcement de la coopération Sud-Sud, à l'inclusion des « nouveaux acteurs » du développement, ou encore à la complémentarité entre l'action des donateurs émergents et des donateurs traditionnels.

[995] KIM, Sung-Mi, MAWDSLEY, Emma, SAVAGE, Laura, 2014, *art. cit.* p. 30.
[996] DAC-OECD, 2011, *op. cit.*

Nous pouvons alors nous demander si le Haut Forum de Busan a marqué la fin de la montée en puissance de la culture de l'efficacité et de l'évaluation dans le secteur de l'aide au développement. Nous estimons plutôt que l'influence des donateurs émergents, conjuguée à l'échec ressenti face aux engagements de la Déclaration de Paris, ont conduit à un changement d'approche en matière d'évaluation. Lors du Haut Forum de Busan, les donateurs ont valorisé sa fonction d'instrument d'apprentissage mutuel, de dialogue, d'échange d'expériences, de transparence et de coordination, plus que son aspect de mesure des résultats et de redevabilité. Tous les donateurs se sont accordés sur l'importance de renforcer la transparence des activités d'aide au développement[997]. Les donateurs émergents ayant participé aux panels sur la coopération Sud-Sud ont souligné l'importance de la mise en place de dispositifs de suivi-évaluation solides pour parvenir à une coopération Sud-Sud efficace, à la condition qu'ils soient élaborés par et à la demande des pays partenaires et qu'ils répondent à un objectif d'amélioration de la prévisibilité de l'aide, de partage de l'information entre les donateurs et d'amélioration de la transparence[998]. Des représentants de KOFID, un réseau d'ONG sud-coréennes travaillant sur l'aide au développement, déclarent à ce propos que « l'importance des résultats a été réaffirmée à Busan. Toutefois, les discussions se sont concentrées sur les systèmes de gestion des résultats et non sur les résultats de la baisse de la pauvreté »[999]. C'est donc une vision *a minima* de l'évaluation qui ressort du Haut Forum de Busan, loin de la conception des dix dernières années qui promouvait des méthodes expérimentales d'évaluation d'impact technocratiques. La fonction participative de l'évaluation est également mise en avant.

La faible importance donnée aux résultats de l'évaluation de la Déclaration de Paris durant le Haut Forum de Busan et la place mineure consacrée dans les sections thématiques aux thématiques du suivi et de l'évaluation laissaient présager que ces enjeux seraient absents de la Déclaration finale. Pourtant, l'importance de l'évaluation y est soulignée à plusieurs reprises. Tout d'abord, une référence aux engagements de Paris et d'Accra a été intégrée à la Déclaration sous la pression des pays bénéficiaires de l'aide : « nous nous appuierons sur nos engagements respectifs, tels que définis dans la Déclaration de Paris et le Programme d'Action d'Accra, pour évaluer conjointement les systèmes nationaux en utilisant des outils de diagnostics approuvés par tous. Sur la base des résultats de ces évaluations, les donateurs décideront du degré d'utilisation des systèmes nationaux »[1000]. Les signataires se sont engagés à « intensifier leurs efforts pour obtenir des résultats tangibles et durables. À cette fin, ils devront : mieux gérer,

[997] *Busan High Level Forum on aid Effectiveness, Proceedings,* décembre 2011.
[998] *Ibid*.
[999]LEE, Seong Hoon, HONG Moon Suk, KANG, Hanee, 2012, *op. cit.*
[1000] DAC-OECD, *Partenariat de Busan pour une coopération efficace au service du développement,* 2011.

suivre, évaluer et faire connaître les progrès réalisés, accroître le renforcement des capacités nationales en suivi-évaluation et prendre des initiatives propres à produire un effet de levier sur les résultats en matière de développement »[1001].

Au-delà de ces engagements, la Déclaration de Busan prévoit de mettre en place un nouvel ensemble d'objectifs, d'indicateurs et de cibles pour évaluer les progrès réalisés par les acteurs de la coopération internationale pour le développement : « nous nous engageons à nous mettre d'accord d'ici juin 2012, sur un ensemble sélectif et pertinent d'indicateurs et d'objectifs, à travers lesquels nous suivrons les progrès de manière continue, en soutenant la redevabilité internationale et régionale de la mise en œuvre de nos engagements »[1002]. L'évaluation est donc un enjeu qui reste très présent dans la Déclaration de Busan, malgré la faiblesse des discussions sur ce thème durant les négociations. Aujourd'hui, dix indicateurs et dix objectifs-cibles[1003] à l'horizon 2015, ont été établis par le groupe intérimaire Post Busan[1004], composé des donateurs du CAD-OCDE, de donateurs émergents et de pays bénéficiaires de l'aide ayant signé la Déclaration de Busan. Les deux-tiers des indicateurs sont quantitatifs et ils informent peu sur la qualité de la coopération pour le développement.

Ainsi, si les donateurs affichent une sortie de « l'ère de l'aide au développement », en témoigne les deux seules références à « l'aide » dans la Déclaration finale de Busan, nous restons toujours dans un contexte où la culture de l'évaluation quantitative et de la transparence est prégnante. Cinq des dix objectifs fixés par le groupe intérimaire portent sur des aspects de mesure des résultats, de transparence, de suivi-évaluation et de redevabilité. L'évaluation est vue comme un instrument d'égalitarisation des relations entre donateurs et bénéficiaires avec le projet de systématiser les évaluations mutuelles entre ces deux catégories d'acteurs. La Déclaration de Busan superpose donc les principes chers aux donateurs traditionnels et aux donateurs émergents, formant un document peu cohérent.

La Déclaration de Busan a été ratifiée *in extremis* par la Chine et l'Inde, le troisième jour du Forum à 2h30 du matin. Ces signatures ont permis aux organisateurs de Busan de célébrer

---

[1001] *Ibid.*
[1002] *Ibid.*
[1003] i) La coopération se concentre sur les résultats qui répondent aux priorités des pays en développement ; ii) la société civile opère dans le cadre d'un environnement qui maximise ses contributions au développement ; iii) engagement et contribution au développement du secteur privé ; iv) transparence : l'information sur la coopération pour le développement est accessible au public ; v) la coopération pour le développement est davantage prévisible ; vi) l'aide est inscrite aux budgets qui sont soumis au contrôle parlementaire ; vii) la redevabilité mutuelle entre les acteurs de la coopération pour le développement est renforcée par des évaluations inclusives ; viii) égalité homme-femme et autonomisation des femmes ; ix) institutions efficaces : les systèmes des pays en développement sont renforcés et utilisés ; x) l'aide est non liée.
[1004] De janvier à juin 2012, le Groupe intérimaire post-Busan (GIPB) a réuni les représentants de tous les acteurs qui ont approuvé le Partenariat de Busan afin de parvenir à un consensus multipartite sur les modalités de travail pour le nouveau partenariat mondial.

le succès de l'évènement lors de la cérémonie de fermeture. Toutefois, considérant les tensions qui ont entouré la participation, puis la signature, de ces pays émergents, il est légitime de se demander si le consensus laborieusement dégagé à Busan entre donateurs émergents et donateurs traditionnels se concrétisera en une réelle collaboration.

B. Une influence normative réciproque

Notre observation du Forum de Busan, nous a amené à nous demander si le clivage entre donateurs traditionnels et donateurs émergents est réellement un point de départ pertinent pour comprendre les clivages qui se créent autour de la notion d'efficacité du développement et des approches de l'évaluation qui en découlent. En effet, l'opposition entre donateurs émergents et donateurs traditionnels nous a semblée assez floue et le positionnement de chacun ambivalent.

- **Des divisions qui transcendent l'opposition entre donateurs émergents et donateurs traditionnels**

Au sein du groupe de travail sur l'efficacité de l'aide du CAD-OCDE, la France, les États-Unis et l'Allemagne ont toujours montré leur réticence quant à la mise en place d'un cadre d'objectifs et d'évaluation mutuelle tel que celui de la Déclaration de Paris. Lors du Forum de Busan, les réticences larvées de ces donateurs à appliquer les « bonnes pratiques » de la Déclaration de Paris se sont révélées au grand jour. Rosalind Eyben rapporte, par exemple, les propos d'un représentant d'une agence d'aide au développement allemande : « nous avons toujours été mal à l'aise avec l'efficacité de l'aide. La coopération est un enjeu de partenariat, d'intérêt mutuel. Ce n'est pas du professorat, de la charité, ni de l'aide. Il n'y a pas de rupture entre bénéficiaire et donateur »[1005]. Il ajoute ensuite que la culture allemande de l'aide au développement se sent proche des principes d'appropriation et de responsabilité mutuelle, tous deux présents dans les principes de la coopération Sud-Sud.

Le Japon semble également mal à l'aise avec son rattachement aux principes du CAD-OCDE. À Busan, les représentants japonais ont mis en avant le fait que leur conception du développement relève d'une culture orientale se rapprochant davantage de celle des donateurs

[1005] EYBEN, Rosalind, SAVAGE, Laura, 2013, *art. cit.*

émergents asiatiques que de celle des donateurs européens et américains[1006]. Nous voyons donc que les divisions sur l'efficacité et l'évaluation de l'aide n'opposent pas seulement les donateurs du CAD-OCDE aux donateurs émergents, mais également les donateurs du Comité entre eux. Face à ce constat, Leidrer et Klingebiel font l'hypothèse que « la division future ne sera peut-être pas entre les donateurs traditionnels et les pays émergents, mais apparaîtra plutôt au sein même de l'OCDE entre les États qui ont été enthousiastes à propos de la Déclaration de Paris et ceux qui ont traîné les pieds, la France, les États-Unis et le Japon »[1007].

Nous avons vu que les donateurs émergents sont relativement divisés quant au positionnement à adopter face aux bonnes pratiques édictées par les donateurs du CAD-OCDE. Ces divisions sont clairement apparues lorsque les donateurs ont dû se positionner sur leur intégration au cadre de suivi-évaluation de la Déclaration de Busan. Le Brésil, l'Inde et la Chine ont informé le groupe intérimaire Post Busan qu'ils ne participeraient à la première réunion de suivi qu'en tant qu'observateurs[1008]. Le comité de direction du groupe a fait pression pour qu'ils participent plutôt en tant qu'« observateurs participants ». De plus, le Brésil et le Mexique ont indiqué qu'étant encore des pays en développement, ils n'étaient pas en mesure d'être associés au cadre de suivi-évaluation[1009]. La Corée du Sud, quant à elle, en tant que membre du CAD, ne pouvait que s'engager pleinement dans ce dispositif. Au final, le Brésil et l'Inde ont assisté à la première réunion du groupe intérimaire post Busan en tant qu'« observateurs participants », mais la Chine n'était pas représentée, marquant ainsi son opposition à la mise en place d'un tel cadre de suivi-évaluation[1010]. Le Mexique, pays hôte, a participé pleinement, mais a mis l'accent sur « le statut particulier des pays qui sont à la fois bénéficiaires et donateurs »[1011]. Face à ces réticences, les donateurs traditionnels ont accepté de reconnaître que les donateurs émergents forment un groupe particulier qui mérite d'être représenté par un co-président au sein du comité de direction du Partenariat Global (qui a succédé au groupe intérimaire), les deux autres groupes représentés étant les donateurs « développés » et les pays bénéficiaires de l'aide. Il a également été décidé que les mêmes progrès ne seraient pas attendus des donateurs émergents et des donateurs traditionnels. L'accord de Busan est donc un engagement à plusieurs vitesses, en termes de redevabilité et de résultats.

[1006] Dr. Akio Hosono, Director, Japan International Co-operation Agency Research Institute, Busan High Level Forum, Table ronde sur l'émergence d'approches asiatiques du développement.
[1007] http://www.thebrokeronline.eu/Blogs/Busan-High-Level-Forum, consulté le 20 décembre 2011.
[1008] OECD-DAC WP EFF, 2012(a), *op. cit.*
[1009] KIM, Sung-Mi, MAWDSLEY, Emma, SAVAGE, Laura, 2014, *art. cit.*
[1010] *Ibid.*
[1011] OECD-DAC WP EFF, 2012(b), *op. cit.*

Nous avons déterminé que ce rejet de l'évaluation provient d'un certain nombre de caractéristiques communes à ces donateurs, que nous allons désormais analyser.

Tout d'abord, les donateurs s'identifient aux pays bénéficiaires en se présentant comme des anciennes victimes de l'exploitation coloniale, comme des pays en (ou ayant été en) développement, ou encore comme les perdants de la mondialisation. Ils jouent de cette identification aux pays bénéficiaires, ou de leur appartenance à la même région géographique, pour créer un sentiment de connivence avec eux. Ils ont une image très technocratique de l'évaluation, servant davantage les intérêts de redevabilité des donateurs et à exercer un contrôle sur les pays bénéficiaires de l'aide, qu'à renforcer le dialogue avec ces derniers et à les guider vers une plus grande autonomie[1012]. Les donateurs émergents rejettent également toute relation hiérarchique entre les États et soulignent les principes de respect de la souveraineté des pays bénéficiaires, d'apprentissage mutuel et d'autonomisation. Enfin, leur approche de la coopération se fonde sur le principe dit « gagnant-gagnant ». Ils veulent ainsi se détacher de la « charité » des donateurs traditionnels qui créerait de la dépendance et ils assument le fait de chercher à tirer profit d'une coopération équitable permettant un bénéfice mutuel entre donateur et bénéficiaire[1013].

La théorie du don de Marcel Mauss permet d'éclairer cette philosophie des pays émergents sur la coopération pour le développement. Dans son ouvrage Essai sur le don[1014], Marcel Mauss analyse la culture du don chez les peuples indigènes du Pacifique, spécifiquement en Polynésie. Il explique que derrière le don qui semble gratuit, il existe une obligation de « contre-don » pour les receveurs, ce qui leur permet de maintenir leur statut et leur honneur et de ne pas rentrer dans un rapport hiérarchique d'infériorité avec les initiateurs du don : « deux éléments du *potlatch* (cérémonie du don) proprement dit sont nettement attestés : celui de l'honneur, du prestige, du *mana* (l'essence spirituelle des hommes et des choses) que confère la richesse, et celui de l'obligation absolue de rendre ces dons sous peine de perdre ce *mana*, cette autorité, ce talisman, cette source de richesse qu'est l'autorité elle-même »[1015]. Marcel Mauss ajoute ensuite : « le don non-rendu rend encore inférieur celui qui l'a accepté, surtout quand il est reçu sans esprit de retour (...) La charité est encore plus blessante

[1012] Constats issus de discussions informelles avec des représentants ministériels laotiens, marocains et ivoiriens rencontrés dans le cadre d'évaluation que nous avons menées pour l'AFD, le ministère des affaires étrangères français et la Cour des comptes française en 2012, 2013 et 2014.
[1013] MAWDSLEY, 2012, *op. cit. Cf.* Chapitre 2.
[1014] MAUSS, Marcel, *Essai sur le don, forme et raison de l'échange dans les sociétés archaïques,* Texte extrait de l'Année sociologique, seconde série, 1924-1925, Éditions PUF, 2007, 248p.
[1015] *Ibid.*

pour celui qui l'accepte, et tout l'effort de notre morale tend à supprimer le patronage inconscient et injurieux du riche aumônier »[1016].

Dans le monde occidental, l'aide est considérée comme un don[1017] dans le sens où les États donateurs n'attendent et ne demandent pas une réciprocité dans l'échange. Les pays bénéficiaires n'ont pas, au moins en apparence, à rendre l'aide qu'ils ont reçue aux pays donateurs. Dans les discours des donateurs émergents, nous retrouvons l'idée selon laquelle cette charité gratuite, ce don unilatéral, crée une relation de pouvoir inégalitaire entre les donateurs et les bénéficiaires de l'aide qui permettrait aux premiers d'affirmer leur supériorité sur les seconds. L'aide charitable serait l'euphémisme d'une domination symbolique du Nord sur le Sud[1018]. Dans le secteur de la coopération internationale, la relation hiérarchisée entre les donateurs et les bénéficiaires créerait une « dépendance à l'aide » de ces derniers[1019]. Les États et leurs populations se mettraient à compter sur l'aide davantage que sur d'autres ressources (impôts, salaires, croissance économique) pour assurer leur développement. Ils auraient alors intérêt à complaire aux conditionnalités des donateurs pour maintenir le flux d'aide qu'ils reçoivent.

Les donateurs émergents voient dans la « charité » des donateurs occidentaux un moyen de maintenir leur supériorité sur les pays du Sud tout en diffusant une image de nations généreuses. Toutefois, cette lecture de l'aide au développement occidentale est caricaturale. En effet, si les donateurs du CAD-OCDE ne lient, en principe, pas directement leur aide à des « contre-dons » économiques ou diplomatiques, ces liens existent implicitement, par exemple avec les idées de sélectivité et de conditionnalité de l'aide. La principale différence est que les donateurs émergents affichent cette idée de « contre-don » et le fait d'attendre en retour de leur aide des privilèges ou des facilités, en termes notamment d'accès aux marchés, d'exploitation des ressources naturelles, etc.

À partir de cette analyse, nous avons réalisé un tableau mettant en relief, de manière idéale-typique, les divergences et les points communs entre l'approche de l'aide des donateurs traditionnels et celle des donateurs émergents[1020] :

[1016] *Ibid*.

[1017] Nous faisons ici référence à l'aide sous forme de don ou de prêts très concessionnels à destination des pays les moins avancés.

[1018] HATTORI, Tomohisa, "The Moral Politics of Foreign Aid", *Review of International Studies,* vol. 29, n°2, avril 2003, pp. 229-247.

[1019] Pour des références plus approfondies sur la théorie du « sentier de dépendance » (path dependency) appliquée à l'aide au développement voir : CHENER, Hollis. B, STROUT Alan, "Foreign Assistance and Economic Development", *American Economic Review*, n°56, 1966, pp. 680-733; LECOMTE, Bernard, NAUDET, Jean-David, 2000, *op. cit.* pp. 177-194.

[1020] Ce tableau est une modélisation et représente les caractéristiques les plus fréquemment retrouvées dans les pratiques d'aide des donateurs composant chacun de ces groupes. Il existe, toutefois, des exceptions à ce modèle.

Tableau 1 : Approches de l'aide des donateurs émergents et occidentaux

Donateurs occidentaux	Donateurs émergents
charité	échange gagnant-gagnant
expertise fondée sur un savoir technique supérieur en termes de gestion des politiques publiques et des systèmes économiques	expertise fondée sur une expérience vécue du développement à partir d'un état de pauvreté
gestion axée sur les résultats, redevabilité mutuelle, alignement, harmonisation, appropriation	égalité, appropriation, respect mutuel, alignement
contrôle de la bonne gouvernance et des politiques économiques et sociales des pays bénéficiaires	non-ingérence dans les politiques des pays bénéficiaires
concentration sur la baisse de la pauvreté et les secteurs sociaux	concentration sur la croissance économique, la production de richesse via l'appui aux investissements, à l'industrie et à la construction d'infrastructure
logique de programmes	logique de projets
transparence : partage de l'information quantitative et qualitative sur l'aide au développement	transparence : partage de l'information quantitative sur l'aide au développement

Source : auteure

De ce tableau, nous déduisons que pour les donateurs émergents l'efficacité d'une activité d'aide s'évalue notamment à partir des bénéfices, diplomatiques ou économiques, qu'elle a entraînés pour ces derniers. Les donateurs du CAD-OCDE, quant à eux, ne prennent pas en compte ce facteur de « retour sur investissement » dans leur évaluation de l'efficacité de l'aide. D'autre part, les donateurs émergents rejettent tout principe « d'ingérence » dans les politiques nationales des pays bénéficiaires et se refusent à utiliser l'évaluation comme un outil de « contrôle social » sur les populations. Les évaluations effectuées par les donateurs émergents se restreignent à l'analyse de la mise en œuvre et des réalisations des activités de coopération[1021]. Les donateurs traditionnels, au contraire, ont une approche de l'évaluation qui se concentre sur la mesure des résultats, des impacts et de la durabilité[1022] de l'aide au développement.

Dans la logique des donateurs émergents, l'évaluation *ex post* de l'efficacité des projets d'aide n'est concevable que si les gouvernements bénéficiaires en formulent la demande. Les acteurs que nous avons pu interroger voient donc ce type d'évaluation comme ne pouvant s'effectuer que de manière mutuelle et à l'initiative des pays bénéficiaires[1023]. Or, comme nous l'avons vu, la demande d'évaluation est encore très faible dans les pays du Sud, ce qui peut expliquer, en partie, pourquoi les donateurs émergents ne pratiquent que très peu d'évaluations *ex post*. Une autre explication, plus convaincante, est que la culture de l'évaluation des

[1021] Entretien avec l'Executive Director, National Center for Science and Technology Evaluation (Chine), 8 juin 2013.
[1022] Les termes réalisations, résultats, impact et durabilité sont définis dans le lexique en annexe 7.
[1023]Entretien avec l'Executive Director, National Center for Science and Technology Evaluation (Chine), 8 juin 2013 ; discussions informelles avec des évaluateurs chinois lors de notre observation participante à l'AFD, durant laquelle l'unité d'évaluation de l'AFD a organisé un dialogue avec des membres chinois du MOFCOM sur l'évaluation de l'aide au développement.

politiques publiques est peu développée dans la plupart des pays émergents[1024]. Les organismes chargés de la coopération pour le développement ne sont donc pas encouragés par leurs tutelles nationales à pratiquer l'évaluation. Les principaux exercices d'évaluation *ex post* auxquels se prêtent les donateurs émergents sont les évaluations conjointes avec les donateurs traditionnels, ces derniers voyant dans ces exercices un moyen de sensibiliser leurs partenaires à l'utilité de l'évaluation[1025].

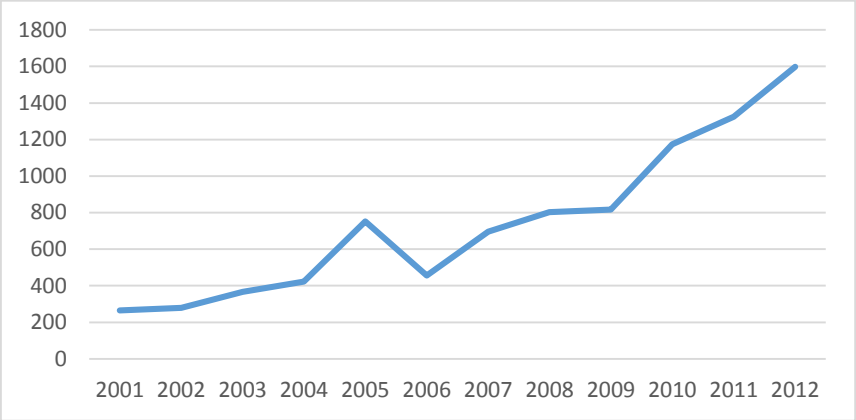
Malgré ces caractéristiques communes, chaque donateur émergent met en place une politique et des pratiques de coopération pour le développement qui lui sont propres, notamment en termes de renforcement de l'évaluation. Nous allons analyser les stratégies divergentes que la Chine et la Corée du Sud ont adoptées pour se créer une identité en tant que donateur et gagner de l'influence dans le régime international de l'aide. Nous analyserons les enjeux politiques qui se nouent autour de l'intégration ou du rejet de la fonction d'évaluation dans ces deux pays et les effets que cela entraîne sur leurs relations avec leurs partenaires internationaux.

- **La Corée du Sud : une appropriation difficile des « bonnes pratiques » d'évaluation**

La Corée du Sud a commencé à développer des activités de coopération internationale pour le développement dès le début des années 1990. Aujourd'hui, deux organismes sont chargés de cette coopération : le Fonds de Coopération pour le Développement Économique (EDCF), fondé en 1987, et l'Agence coréenne de Coopération internationale pour le développement (KOICA), créée en 1991[1026]. La Corée du Sud ne rapporte ses données d'aide au développement au CAD-OCDE que depuis 2007. Depuis cette date, l'aide au développement du pays a plus que doublé, passant de 696 millions de dollars en 2007 à 1, 6 milliards de dollars en 2012[1027].

[1024] FURUBO, Jan-Eric, RIST, Ray. C, SANDAHL, Rolf, 2002, *op. cit.*
[1025] Discussions informelles avec des évaluateurs chinois lors de notre observation participante à l'AFD, durant laquelle l'unité d'évaluation de l'AFD a organisé un dialogue avec des membres chinois du MOFCOM sur l'évaluation de l'aide au développement.
[1026] CHUN, Hong Mi, MUNYI, Elijah N, LEE, Heejin, "South Korea as an Emerging donor: challenges and change on its entering OECD-DAC", *Journal of International Development,* vol. 22, 2010, pp. 788-802.
[1027] DAC-OECD QWIDS, https://stats.oecd.org/qwids/#?x=1&y=6&f=4:1,2:1,3:51,5:3,7:1&q=4:1+2:1+3:51+5:3+7:1+1:61+6:2007,2008,2009,2010,2011,2012, consulté le 21 mars 2014.

Graphique 4 : Évolution de l'aide sud-coréenne de 2001 à 2012



Source: International Development Statistics Online DB (OECD) ; Korea Eximbank.

Le graphique ci-dessus montre une forte croissance de l'aide sud-coréenne à la suite de sa décision d'adhérer au CAD-OCDE en 2009.

La Corée du Sud a connu un développement extrêmement rapide et efficace durant les cinquante dernières années[1028]. Fort de ce succès, le gouvernement sud-coréen entend proposer aux pays en développement d'adopter le modèle sud-coréen de développement ; une version moderne de la théorie « l'état développemental »[1029]. La Corée du Sud joue donc sur l'unicité et le caractère exceptionnel de son expérience de développement pour légitimer son statut de donateur[1030]. Ce succès de la Corée du Sud est unanimement reconnu. Beaucoup d'études ont été réalisées pour analyser le phénomène du développement sud-coréen et comprendre les clefs de cette réussite[1031].

La Corée du Sud cherche à créer un sentiment de confiance avec les bénéficiaires de l'aide en mettant en avant le fait qu'elle a été, pendant plusieurs décennies, un pays fortement soutenu par l'aide internationale. Durant cette période, le gouvernement sud-coréen a tenté de garder le contrôle sur la gestion de ses politiques publiques, malgré la pugnacité des donateurs occidentaux à vouloir imposer leurs modèles économiques[1032]. De là, provient l'attachement de

[1028] . Au sortir de la guerre de Corée en 1953, la Corée du Sud était un État dévasté et ruiné, bien plus pauvre que son voisin la République démocratique de Corée du Nord soutenue par l'URSS. Partant de cette situation de développement désastreuse, la Corée du Sud était le 15ème PIB mondial en 2012, et au 12ème rang mondial en termes d'IDH.

[1029] Le modèle du "développemental State" réfère au phénomène de planification macroéconomique encadrée par l'État dans plusieurs États de l'Asie de l'Est au XXe siècle. Dans ce modèle capitaliste, l'État a un pouvoir fort sur l'économie qu'il contrôle étroitement. L'État se veut autonome et interventionniste. Pour plus de détail sur ce modèle de développement, voir : LEFTWICH, Adrian, *States of Development : on the Primacy of Politics in Development*, Polity Press, 2000, 240p ; AMSDEN, Alice, H. *Asias's Next Giant : South Korea and late Industrialization*, Oxford University Press, 1992, 400p.

[1030] KIM, Eun Mee, KIM, Pil Ho, KIM Jin Kyung, "From development to development cooperation: foreign aid, country ownership and the developmental state in South Korea", *The Pacific Review*, vol. 26, n°3, février 2013, pp. 313-336.

[1031] Par exemple, KOICA-UNRISD, *Making International Development Cooperation Effective*, 2011.

[1032] MACDONALD, Donald, *US-Korea relations form liberation to Self-Reliance: the twenty-year record*, Westview PR, 1992, 365p.

la Corée du Sud à préserver l'autonomie des pays qui reçoivent aujourd'hui son aide au développement.

Ainsi, un dirigeant de l'agence d'aide KOICA a déclaré dans un entretien mené en 2014 : « *la Koica cherche à diffuser une version modernisée du modèle du Mouvement Saemul*[1033] *à travers le monde* »[1034]. Selon le gouvernement sud-coréen, un processus de développement réussi dépend « d'une combinaison fortuite d'appropriation nationale des bonnes pratiques proposées par les donateurs, d'autonomie vis-à-vis des puissances étrangères et de renforcement des capacités »[1035]. Nous retrouvons ces principes dans les « bonnes pratiques » de l'aide au développement sud-coréenne : le respect de la souveraineté des pays en développement et de leurs choix de politiques publiques, l'alignement sur les programmes de développement nationaux, l'appropriation et l'autonomisation[1036].

La Corée du Sud, pour « vendre » son modèle de développement, joue sur le fait qu'il est accessible aux pays du Sud, puisque reposant en grande partie sur l'aide au développement reçue dans le passé[1037]. La Corée du Sud cherche à se distinguer des autres donateurs par sa *succes story* en tant que pays bénéficiaire d'aide. Ainsi, le ministère des affaires étrangères déclarait en 2013 : « nous travaillons à créer un modèle coréen unique d'aide au développement se concentrant sur les secteurs dans lesquels nous avons un avantage comparatif par rapport aux autres États, et notamment par rapport aux membres du CAD-OCDE »[1038].

Nous pouvons alors nous demander pourquoi la Corée du Sud a décidé d'adhérer au CAD-OCDE. En effet, cette décision semble contradictoire avec le discours du gouvernement sud-coréen qui prône un modèle unique de coopération pour le développement, éloigné de celui des donateurs traditionnels. Rejoindre le CAD-OCDE signifiait pour la Corée du Sud rejoindre le « clan » des donateurs traditionnels et perdre son identité de donateur offrant une alternative par rapport aux modèles prônés par les membres du CAD-OCDE. La Corée du Sud allait devoir adhérer à l'agenda sur l'efficacité de l'aide et à ses « bonnes pratiques », notamment en termes de gestion axée sur les résultats, de transparence, de redevabilité mutuelle et d'évaluation.

[1033] Nom donné au modèle coréen de développement économique tiré du plan national de développement rural mis en place par le gouvernement coréen dans les années 1970-1980.
[1034] Entretien avec un dirigeant de la KOICA, Séoul, le 26 mai 2014.
[1035] MARX, Axel, SOARES, Jadir, "South Korea's Transition from Recipient to DAC Donor: Assessing Korea's Development Cooperation Policy", *International Development Policy*, 2013, pp. 107-142.
[1036] KIM, Eun Mee, KIM, Pil Ho, KIM Jin Kyung, 2013, art. cit.
[1037] MARX, Axel, SOARES, Jadir, 2013, art. cit.
[1038] Ministry of Foreign Affairs, Republic of Korea, ODA/Development Cooperation page http://www.mofat.go.kr/ENG/policy/oda/index.jsp?menu=m_20_11, consulté le 3 avril 2013.

D'un autre côté, son adhésion au CAD-OCDE allait donner un signal sur la scène internationale de la puissance sud-coréenne. Le gouvernement sud-coréen pouvait aussi espérer gagner en influence au sein du régime de l'aide, et ainsi, avoir plus de poids par rapport à la Chine dans sa stratégie de conquête de marchés et de ressources dans les pays en développement[1039]. Rejoindre le CAD permettait à la Corée du Sud de gagner une réputation de donateur qui respecte les bonnes pratiques du régime international de l'aide, contrairement à la Chine ou à l'Inde. Nous constatons, en effet, que, dès son entrée au CAD-OCDE, la Corée du Sud est symboliquement passée du statut de donateur émergent à celui de donateur traditionnel et a donc été « reclassée » sur la scène internationale. Elle est vue comme un « bon » donateur et n'est plus évoquée comme étant un donateur émergent dans la littérature et dans les discours des donateurs, bien que son activité d'aide soit plus récente et moins importante que celle de la Chine.

Nous allons désormais étudier dans quelle mesure l'adhésion de la Corée du Sud au CAD a fait évoluer ses pratiques d'aide, et plus spécifiquement d'évaluation.

En adhérant au CAD-OCDE, la Corée du Sud s'engagea à réformer ses pratiques d'aide pour les faire correspondre aux standards de l'organisation. En 2009, le gouvernement sud-coréen a annoncé sa volonté d'augmenter le ratio d'APD par rapport au PIB de 0, 09 % à 0,25 % d'ici 2015. Il a également promis de faire des efforts pour améliorer la planification, l'allocation et l'évaluation de ses activités d'aide. Enfin, il s'est engagé à « systématiquement documenter les succès et les erreurs expérimentés durant les activités de coopération pour le développement »[1040]. En 2010, le ministère des affaires étrangères sud-coréen affirmait que « les principes et les bonnes pratiques de l'efficacité de l'aide contenus dans les déclarations de Paris et d'Accra sont des piliers de la politique de coopération sud-coréenne pour le développement »[1041]. Le défi pour la Corée du Sud est désormais de mettre en pratique ces engagements. Il est particulièrement important en ce qui concerne les aspects de gestion axée sur les résultats et d'évaluation qui n'étaient, jusqu'alors, pas développés dans la culture organisationnelle des organismes de développement.

Une des premières réformes entreprises par le ministère des affaires étrangères sud-coréen a été de donner un cadre d'objectifs à la coopération pour le développement. L'enjeu était de préserver les principes sud-coréens d'aide au développement tout en se conformant

[1039] CHUN, Hong Mi, MUNYI, Elijah. N, LEE, Heejin, 2013, *op. cit.*
[1040] OECD-DAC, *Korea: Peer Review 2012*, 2012.
[1041] *Ibid.*

aux bonnes pratiques du CAD-OCDE. Ces objectifs sont : augmenter le volume d'aide et la part d'aide sous forme de don ; i) mieux « sélectionner » les pays bénéficiaires et les secteurs de l'aide ; ii) mettre en place un système de gestion axée sur les résultats qui répond aux besoins des pays partenaires ; iii) intégrer l'expérience de développement sud-coréenne et ses avantages comparatifs dans les projets et programmes ; iv) délier l'aide : v) encourager l'implication du secteur privé et de la société civile dans les projets et programmes ; vi) renforcer les capacités de mise en place, d'évaluation et de de l'aide pour améliorer son efficacité[1042]. Certains de ces objectifs reflètent clairement les exigences du CAD-OCDE en termes de gestion de l'aide au développement (i ; ii ; iv ; vi) et sont parfois contraire aux pratiques habituelles de l'aide sud-coréenne. En effet, traditionnellement, l'aide est liée aux intérêts économiques nationaux et est ciblée sur les pays qui offrent les meilleures perspectives de résultats et « opportunités » pour la Corée du Sud, en termes de ressources naturelles, de développement de nouveaux marchés, ou de diplomatie. D'autres objectifs reflètent davantage l'approche coréenne de l'aide au développement (iii ; v). Nous remarquons que deux des objectifs portent sur le renforcement d'une culture de l'évaluation, ce qui témoigne de la pression exercée par le CAD-OCDE pour que la Corée du Sud augmente ses capacités sur ce plan. En conséquence, les organismes d'aide sud-coréens ont dû développer, en quelques années, des systèmes opérationnels d'évaluation.

L'entrée de la Corée du Sud au sein du CAD-OCDE s'est, en effet, accompagnée d'un effort d'apprentissage en termes d'évaluation. Il apparaît clairement que les acteurs sud-coréens de l'aide au développement souhaitent se nourrir de l'expérience des donateurs plus expérimentés et s'aligner, au moins formellement, sur leurs modèles[1043]. Les représentants du CAD-OCDE que nous avons pu interroger déclarent que les Sud-coréens sont « très assidus aux réunions du réseau d'évaluateurs du CAD-OCDE » et qu'ils « veulent apprendre, même si le schéma d'évaluation du CAD-OCDE n'est pas adapté à leur fonctionnement »[1044]. Les dirigeants de la KOICA, quant à eux, font fièrement part des nombreux accords d'échanges d'expérience qu'ils ont conclus avec les États-Unis, la France ou encore le DFID[1045].

Nous observons une corrélation entre la décision du gouvernement sud-coréen de candidater à l'adhésion au CAD-OCDE en 2006 et la création d'une unité d'évaluation au sein

[1042] Voir le site de la Koica qui présente les objectifs de la coopération internationale sud-coréenne : http://www.koica.go.kr/english/koica/policies/index.html, consulté le 11 avril 2014.
[1043] Ce constat se fonde sur notre participation à un programme d'échange organisé en 2011-2012 encadré par le Korean Institute for International Economic Policy qui visait à renforcer les compétences des experts coréens en matière de gestion et de suivi-évaluation des activités d'aide au développement et de développer la coopération entre les experts d'Europe et de Corée du Sud.
[1044] Entretiens au Siège du CAD-OCDE.
[1045] Entretien avec un évaluateur de la KOICA, 26 juin 2014 ; entretien avec un chargé de projets sur l'efficacité de l'aide de la KOICA, 26 juin 2014.

de l'EDCF en janvier 2007. La KOICA, quant à elle, possédait une unité d'évaluation depuis 1996. Cette unité d'évaluation était située dans le département de la planification et n'était donc pas séparée des départements opérationnels[1046]. Cette pratique allait à l'encontre des standards de l'OCDE qui recommandent l'indépendance des structures d'évaluation. En conséquence, l'unité d'évaluation a été réorganisée en 2006. Dans un premier temps, il a été décidé qu'elle rendrait des comptes au directeur exécutif de l'agence, puis en 2011, à son président directement pour encore plus d'indépendance.

De plus, un sous-comité pour les évaluations a été créé, afin « *d'établir une politique d'évaluation de l'aide au développement, contrôler la qualité des évaluations menées par la KOICA et l'EDCF, établir des plans annuels d'évaluation avec les organismes d'aide et assurer le suivi des recommandations des évaluations* »[1047]. Ce sous-comité, placé sous l'égide du cabinet du premier ministre, est composé de quatre représentants de l'EDCF et de la KOICA et de 11 experts indépendants. Il a été créé sur la recommandation du CAD-OCDE pour améliorer la qualité des évaluations et pour augmenter leur utilisation[1048]. Avec l'aide du CAD-OCDE, le sous-comité a établi des « bonnes pratiques » d'évaluation à destination de tous les ministères et agences impliqués dans l'aide au développement. De plus, il commissionne annuellement des évaluations sur des enjeux et des thématiques qui ont un intérêt stratégique particulier pour les dirigeants sud-coréens. Tous les acteurs de l'aide sont requis de rendre compte au sous-comité de la manière dont ils entendent mettre en œuvre les recommandations des évaluations.

À ce suivi des évaluations, s'ajoutent des mécanismes internes de « feedback » qu'ont mis en place la KOICA et l'EDCF à la demande du CAD-OCDE. Une base de données a récemment été créée, afin d'assurer le suivi des projets et programmes et ainsi de faciliter les exercices d'évaluation[1049]. La Corée du Sud a donc mis en œuvre des réformes structurelles significatives en matière d'évaluation pour se conformer aux bonnes pratiques du CAD. À tel point que ses structures d'évaluation paraissent, sur le papier, plus sophistiquées que celles de certains donateurs traditionnels, la France par exemple.

Au-delà de ces réformes structurelles, l'activité d'évaluation s'est largement développée dans les agences d'aide sud-coréennes. Alors que de 1998 à 2004, la KOICA réalisait en moyenne 8,5 évaluations par an, depuis 2005, elle en réalise environ 16,5[1050]. Le nombre

[1046] Entretien avec un chargé de projets sur l'efficacité de l'aide au sein de la KOICA, 26 juin 2014.
[1047] Entretien avec un évaluateur de la KOICA, 26 juin 2014.
[1048] OECD-DAC, 2012, *op. cit.*
[1049] Entretien avec un ancien membre de la KOICA, 27 juin 2014.
[1050] Entretien avec un évaluateur de la KOICA, 26 juin 2014.

d'évaluations a donc doublé avec la perspective d'entrée, puis l'adhésion, au CAD-OCDE. Le budget que la Corée du Sud consacre à l'évaluation est d'environ 800 000 dollars par an, soit 0,22 % du budget total de l'aide, ce qui est deux fois plus élevé que la moyenne des membres du CAD[1051].

La transparence des évaluations a également été améliorée, dans la mesure où les rapports d'évaluations, ou leurs conclusions, sont publiés sur les sites Internet des organismes d'aide[1052]. L'effort de la Corée du Sud pour augmenter la qualité de ses évaluations, et ainsi se hisser au niveau des autres donateurs, transparaît également dans les initiatives qu'elle a menées en matière de renforcement des capacités en évaluation de son personnel. Par exemple, en 2011, la KOICA a organisé des formations sur la gestion axée sur les résultats et le cadre logique menées par des experts japonais[1053]. Les évaluateurs de la KOICA se félicitent des avancées faites par l'organisation en matière d'évaluation : « *pour ce qui est de la mesure de l'efficacité de notre APD, nous avons réussi à mettre en place un système d'évaluation poussé* »[1054].

Malgré ces efforts, le CAD-OCDE exerce encore une pression forte sur la Corée du Sud pour qu'elle améliore davantage ses capacités d'évaluation. Le CAD-OCDE encourage la décentralisation de la fonction d'évaluation vers les agences de terrain, comme c'est la norme chez les donateurs traditionnels. Les organismes d'aide sud-coréens sont également appelés à mener davantage d'évaluations d'impact et conjointes. Sous cette pression normative, la Corée du Sud se doit donc de faire évoluer très rapidement ses pratiques d'évaluation en reproduisant le modèle du CAD-OCDE. Peu de place est laissée à l'adaptation et à l'appropriation.

En conséquence, nous retrouvons dans le système d'évaluation de l'aide sud-coréenne un paradoxe déjà identifié plusieurs fois dans notre recherche. Alors que des structures sophistiquées d'évaluation se mettent en place et que le nombre d'évaluations augmente, la Corée du Sud éprouve des difficultés à créer une expertise en évaluation suffisamment qualifiée et conséquente pour assurer la production d'évaluations répondant aux critères de qualité du CAD-OCDE. Pour solutionner ce problème, le sous-comité des évaluations a commandité une évaluation du système d'évaluation de l'aide au développement sud-coréenne, afin d'identifier des voies d'amélioration[1055]. Avant même d'être complètement abouti et éprouvé, ce système d'évaluation est donc déjà soumis à l'évaluation. L'évaluation devient aujourd'hui la première

[1051] KIM *et al*, 2011, *op. cit.*
[1052] EDCF : http://www.edcfkorea.go.kr/edcfeng/bbs/pubanew/list.jsp?bbs_code_id=1368577018458&bbs_code_tp=BBS_4, KOICA : http://www.koica.go.kr/english/schemes/evaluation/report/index.html, consultés le 11 avril 2014.
[1053] Entretien avec un ancien membre de la KOICA, 27 juin 2014.
[1054] Entretien avec un évaluateur à la KOICA, 26 juin 2014.
[1055] Entretien avec un ancien membre de la KOICA, 27 juin 2014.

réaction des acteurs de l'aide, lorsqu'un problème ou un dysfonctionnement est observé. Nous assistons à une surenchère de l'évaluation dans le secteur de l'aide au développement.

Si, en apparence, « *la politique, le système, les ressources, les procédures et les performances d'évaluation de la Corée du Sud peuvent être jugés satisfaisants, dans la mesure où ils sont comparables à ceux des donateurs traditionnels* »[1056], lorsque nous analysons plus en profondeur l'inscription des pratiques d'évaluation dans les cultures organisationnelles des agences d'aide sud-coréennes, il apparaît que les évaluations sont semblables à des « ilots d'excellence » déconnectés du reste de la gestion de l'aide qui reste inchangé. L'évaluation est peu utilisée et peu intégrée aux processus de mise en œuvre des programmes d'aide et à la prise de décision[1057].

L'évaluation n'est pas perçue comme un outil d'apprentissage et d'amélioration des pratiques, mais comme un instrument de sanction. Certains voient une explication culturelle à cette vision de l'évaluation en Corée du Sud : « *dans la culture sud-coréenne, l'évaluation et les critiques qu'elle peut entraîner tendent à être associées au blâme et les évalués ont alors l'impression de « perdre la face » ou d'être dévalorisés* »[1058]. Nous estimons plutôt que ces blocages ont les mêmes causes que dans les agences d'aide occidentales : une pression à décaisser rapidement des fonds qui laisse peu de temps à l'apprentissage, le manque d'incitation à l'utilisation des évaluations, une forte croyance du personnel des agences d'aide dans le bien-fondé de leurs pratiques, une faible mémoire institutionnelle et des objectifs flous.

En intégrant le CAD-OCDE, la Corée du Sud s'estime stratégiquement bien placée pour « changer de l'intérieur » le régime de l'aide au développement, en défendant les intérêts des donateurs émergents et en faisant du plaidoyer pour la reconnaissance de l'utilité et de l'efficacité de la coopération Sud-Sud[1059]. Son ambition est de gagner une réputation de donateur efficace, respectant les bonnes pratiques internationales de l'aide au développement, tout en gardant son image de donateur « unique » et proche des pays en développement. Cette stratégie est risquée, dans la mesure où la pression normative du CAD-OCDE est telle que la Corée du Sud se retrouve contrainte à appliquer, au moins en partie, les principes du CAD, ce qui entraîne un décalage de plus en plus flagrant entre, d'une part les pratiques et, d'autre part les discours des acteurs de l'aide sud-coréens. Les bonnes pratiques du CAD-OCDE se trouvent

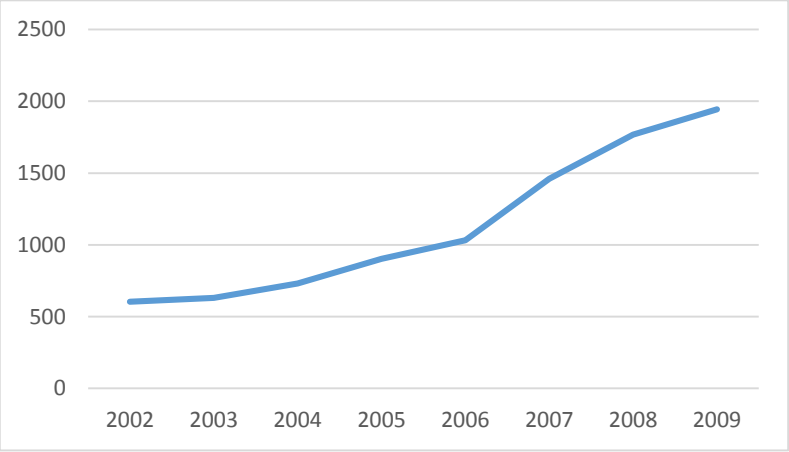
[1056] *Ibid*.
[1057] DAC-OECD, 2012, *op. cit.*
[1058] Entretien avec un ancien membre de la KOICA, 27 juin 2014.
[1059] PARK, Kang Ho, "Korea's Role in Global Development", *Brookings*, Février 2010 : http://www.brookings.edu/research/opinions/2010/02/09-korea-global-development, consulté le 24 mars 2014.

superposées aux anciens principes de l'aide sud-coréenne formant un ensemble peu cohérent, avec, par exemple, d'un côté la proclamation du rejet des pratiques de sélectivité des donateurs traditionnels et de l'autre des engagements auprès du CAD-OCDE à mieux sélectionner les pays bénéficiaires selon des critères de pauvreté et de bonne gouvernance. Toutefois, nous pouvons penser qu'avec l'abandon de l'agenda sur l'efficacité de l'aide et l'adoption du partenariat de Busan, dont les principes sont plus en accord avec ceux de la Corée du Sud, ce déséquilibre tendra à s'atténuer.

- **La Chine : une politique d'aide poreuse aux bonnes pratiques des donateurs occidentaux**

Comme nous l'avons vu au début de ce chapitre, la Chine développe des activités d'aide au développement[1060] depuis les années 1950. Toutefois, ce n'est que depuis le début des années 2000 que ces activités s'étendent intensivement et extensivement en dehors de la région asiatique, spécifiquement en Afrique sub-saharienne. Le graphique ci-dessous montre que l'aide au développement chinoise a plus que triplé de 2002 à 2009 :

Graphique 5 : Évolution de l'aide chinoise de 2002 à 2009 (MUSD)



Sources : tiré des données du tableau « Main items of budgetary expenditure of central government » MOFCOM

Depuis 1964, l'aide au développement chinoise est régie par huit principes : i) la Chine fournit une aide étrangère selon les principes d'égalité et d'avantages mutuels ; ii) lorsqu'elle fournit une aide, la Chine n'impose jamais de conditions ni ne réclame aucun privilège ; iii) la

[1060] Il est toutefois important de noter que l'aide au développement chinoise n'est pas tout à fait comparable à l'APD des bailleurs de fonds du CAD, dans la mesure où l'aide n'est souvent qu'une partie d'un ensemble plus large d'accords d'investissement ou d'échanges commerciaux. L'aide chinoise est considérablement liée à la promotion des intérêts économiques de la Chine dans les pays bénéficiaires.

Chine s'efforce d'alléger du mieux possible le fardeau sur les pays bénéficiaires ; iv) la Chine a pour but d'aider les pays bénéficiaires à devenir progressivement indépendants et autonomes ; v) la Chine s'efforce de développer des projets d'aide qui demandent des investissements peu élevés mais suivis de résultats rapides ; vi) la Chine fournit des équipements et du matériel chinois de la meilleure qualité; vii) en fournissant une assistance technique, la Chine s'engage à faire pleinement bénéficier le personnel du pays bénéficiaire de ce savoir technique ; viii) les experts chinois ne sont pas autorisés à soumettre de demande particulière ou à bénéficier d'aucune faveur[1061]. Ces principes mettent en avant les idées de bénéfice mutuel, d'égalité, de respect de l'indépendance des pays bénéficiaires et d'efficacité. L'aide est perçue comme un instrument de court-terme pour stimuler la croissance économique des pays en développement.

En avril 2011, la Chine a produit son premier *White Paper* sur l'aide au développement. Ce document met l'accent sur le fait que la Chine est, elle aussi, un pays en développement. Ainsi, elle se met sur un pied d'égalité avec les pays bénéficiaires de ses actions de coopération internationale : « la Chine est un pays en développement. Au fil des ans, alors qu'elle se concentrait sur son propre développement, la Chine a fourni du mieux qu'elle le pouvait de l'aide aux autres pays en développement connaissant des difficultés économiques, remplissant ainsi son devoir international (…) L'aide étrangère chinoise a des caractéristiques distinctes (de celle des autres donateurs). La Chine est le plus grand pays en développement au monde, avec une forte population qui souffre de pauvreté et d'un développement économique inégalitaire. Étant donné que le développement reste une tâche ardue et de long terme pour la Chine, son aide internationale rentre dans la catégorie de la coopération Sud-Sud et promeut l'assistance mutuelle entre les pays en développement »[1062]. La Chine justifie donc ses pratiques d'aide liée[1063] par son statut de pays « pauvre » et met l'accent sur les difficultés communes qu'elle partage avec les pays du Sud. C'est une stratégie pour se distinguer des donateurs occidentaux qui ne comprendraient pas les tourments des pays en développement dans la mondialisation. D'un autre côté, la Chine met en avant ses efforts de dialogue et d'échange avec les autres donateurs : « la Chine a envoyé des délégations participer aux conférences et aux dialogues sur la coopération internationale pour le développement, tels que le Haut Forum des Nations unies sur le Financement du Développement, Le Haut Forum des Nations unies sur les Objectifs du

[1061] ENLAI, Zhou, *Eight Principles for China's Aid to Foreign Countries,* 1964.
[1062] People's Republic of China, 2011a, *op. cit.*
[1063] Pour une définition des concepts « d'aide liée » et « d'aide déliée », voir lexique en annexe 7.

Millénaire, le Forum des Nations unies sur la Coopération pour le Développement et Le Haut Forum sur l'Efficacité de l'aide. »[1064].

Malgré cette apparente ouverture à la discussion, la Chine refuse toujours d'adhérer à un quelconque accord international l'engageant à respecter un ensemble de bonnes pratiques partagé avec les donateurs occidentaux. C'est ainsi qu'à la suite du Haut Forum de Busan, la Chine a refusé de participer à l'élaboration du cadre d'évaluation des engagements adoptés. Les donateurs traditionnels blâment ce refus d'engagement de la Chine, la stigmatisant comme un « mauvais » donateur, peu transparent et peu redevable. Toutefois, leurs critiques sont mises à mal lorsque l'on voit la manière dont ils se sont désengagés de la Déclaration de Paris, saisissant l'opportunité ouverte par la Corée du Sud et l'ensemble des donateurs émergents d'enterrer l'agenda sur l'efficacité de l'aide pour le remplacer par le partenariat global pour le développement, plus inclusif mais moins contraignant[1065].

La Chine cherche à afficher son opposition aux principes actuels du régime international de l'aide. C'est une stratégie pour séduire les pays en développement en leur proposant une aide plus souple, moins technocratique et non-conditionnée à des critères de bonne gouvernance. Toutefois, notre analyse nous amène à relativiser les différences entre les bonnes pratiques des donateurs du CAD-OCDE et de la Chine sur l'efficacité de l'aide.

Tableau 2 : Bonnes pratiques du CAD-OCDE / bonnes pratiques de l'aide chinoise

Bonnes pratiques internationales	**Bonnes pratiques chinoises**
Objectifs du Millénaire	Oui
Déclaration de Paris	
Appropriation	Oui
Alignement	Oui
Harmonisation	Non
Résultats	Quantitatif mais pas qualitatif
Redevabilité	Non
Conditionnalité	Pas politique

Source : auteure, inspiré de la thèse de Philippa Brant : *No Strings attached : Chinese Foreign Aid and its Implications for the International Aid Regime,* mai 2012.

[1064] *Ibid.*

[1065] TALAAT, Abdel-Malek, *The Global Partnership for Effective Development Cooperation: What next for a global architecture,* The Current Column, DIE-GDI, 2002; FUES, Thomas, KLINGEBIEL, Stephan, *Unrequited love: What is the legacy of the first Global Partnership summit,* The Current Column, DIE-GDI, 2014.

Ce tableau montre que les principes chinois d'efficacité de l'aide ne sont pas si éloignés qu'il ne le paraît de ceux des donateurs traditionnels. La Chine a adhéré aux OMD et suit, à sa manière, des « bonnes pratiques » d'appropriation par les pays bénéficiaires et d'alignement des projets d'aide sur les priorités nationales. Bien que les OMD ne soient pas la priorité du gouvernement chinois, les officiels du régime y font souvent référence dans leurs discours. Le porte-parole du ministère des affaires étrangères, Hong Lei, a ainsi déclaré en 2011 : « la Chine améliorera la performance de ses prêts et en dédiera davantage aux domaines d'action des OMD dans un effort pour mieux assister le développement économique et social des pays bénéficiaires »[1066]. En revanche, nous ne retrouvons pas dans la politique d'aide chinoise d'exigences en termes de redevabilité et de gestion axée sur les résultats, ce qui implique une pratique de l'évaluation restreinte à une mesure quantitative des allocations d'aide. Enfin, si la Chine se prévaut de ne pas attacher de conditionnalités de bonne gouvernance ou de performance à son aide, sa politique de bénéfices mutuels revient à une forme de conditionnalité économique. En effet, les pays en développement ne peuvent recevoir de fonds d'aide chinois qu'en échange de ressources naturelles, d'ouvertures de marché pour les technologies chinoises et de contrats industriels et économiques.

Toutefois, le concept d'efficacité de l'aide est présent dans les documents politiques du gouvernement chinois. Dans le *White Paper* de la politique d'aide au développement chinoise, il est écrit que « la Chine fera des efforts pour optimiser la structure du système national d'aide au développement, améliorer la qualité de l'aide internationale, renforcer la capacité des pays partenaires à atteindre un développement indépendant, et améliorer la pertinence et l'efficacité de l'aide »[1067]. Nous trouvons une autre référence à l'efficacité de l'aide dans le *White Paper* sur le développement pacifique de septembre 2011 : « nous améliorerons notre système d'allocation de l'aide, pour le rendre plus efficace »[1068]. Par ailleurs, nous constatons que la Chine augmente sa contribution aux banques régionales de développement et à la Banque mondiale[1069]. Or, dans le cadre de ses engagements avec des institutions multilatérales, la Chine doit respecter leurs bonnes pratiques, notamment en termes de suivi-évaluation et de transparence. En intégrant le système de l'aide, la Chine se voit donc obligée de réformer ses pratiques d'aide vers plus d'évaluation des résultats et de transparence.

[1066] Xinhua, "China says soft loans to Pacific island nations well-received", *People's Daily*: http://english.people.com.cn, consulté le 10 avril 2012.
[1067] People's Republic of China, 2011a, *op. cit.*
[1068] People's Republic of China, *China's Peaceful Developmen*t. Beijing, China: Information Office of the State Council, 2011.
[1069] BRÄUTIGAM, Deborah, *China, Africa and the International Aid Architecture,* working paper, Banque africaine de Développement, mars 2010.

De plus, la Chine cherche à se défaire de son image de bailleur ayant des « mauvaises pratiques » qui peut nuire à sa réputation au niveau local et international, d'autant plus que les donateurs traditionnels n'hésitent pas à critiquer la Chine lors des sommets internationaux. Hillary Clinton, par exemple, dénonçait, de manière à peine voilée, les pratiques d'aide chinoises lors de la cérémonie d'ouverture du Haut Forum de Busan : « nous devons tous nous mettre au niveau des standards internationaux auxquels la communauté internationale a adhéré, en commençant par l'engagement d'aider les pays à devenir plus autonomes. Cela signifie, par exemple, aider les pays qui possèdent des ressources naturelles à échapper à la « malédiction des ressources naturelles » qui les laisse riches en pétrole, en or, ou en autres marchandises, mais pauvres en bien d'autres façons. Alors que la souveraineté nationale est un principe important, il ne doit pas devenir une excuse pour éviter l'évaluation approfondie de nos efforts de développement, pas si nous voulons des résultats. La transparence aide à révéler nos faiblesses, afin que nous puissions améliorer notre travail »[1070].

La Chine est sensible aux critiques de ses pairs, ainsi qu'à celles de la société civile internationale qui l'accuse, par exemple, « de s'accaparer des terres en Afrique »[1071]. En conséquence, depuis quelques années, la Chine s'est engagée à améliorer la qualité de son aide au développement, notamment *via* la mise en place d'évaluations. Selon des chercheurs occidentaux et chinois, la réglementation chinoise stipule que les projets doivent être accompagnés d'évaluations *ex ante* et *ex post*. Le problème est que cette règlementation ne précise pas ce que ces évaluations doivent contenir, ni les standards qualitatifs qu'elles doivent respecter[1072].

Depuis quelques années, nous observons des initiatives de dialogue provenant de la Chine, afin d'échanger des bonnes pratiques avec les donateurs occidentaux, notamment sur la coopération en Afrique. Par exemple, en 2010, la Chine a convié à Beijing des donateurs occidentaux à un forum sur le thème de la coopération Chine-Afrique, afin de recueillir l'expérience de ces derniers sur ce terrain de l'aide. Des entretiens nous ont révélé que les participants chinois (universitaires, fonctionnaires, représentants d'entreprises) du forum ont montré une réelle volonté d'améliorer la transparence de leur aide au développement et de mieux communiquer sur les impacts de la présence chinoise sur le développement économique

[1070] RODHAM CLINTON, Hillary, *Speech from the Opening Ceremony of Busan High Level Forum on Aid Effectiveness*, novembre 2011.
[1071] Discussion informelle avec un représentant d'une organisation d'aide au développement occidentale ayant participé au Forum 2010 *Chine-Afrique: commerce, investissement et coopération technique dans le secteur agricole,* 10-12 décembre 2010 à Abidjan, le 25 mai 2014.
[1072] *Ibid*.

des pays africains, sans pour autant remettre en cause la poursuite de leurs intérêts, notamment dans le secteur agro-alimentaire[1073].

Dans sa thèse sur l'aide au développement chinoise, Philippa Brant relate que les officiels chinois ont été consternés au Haut Forum de Busan, lorsqu'il a été acté que le principe d'efficacité de l'aide serait remplacé par celui d'efficacité du développement, car ils avaient travaillé durant des années auprès de leur gouvernement pour que ce terme soit accepté et intégré à la politique de l'État[1074]. Cette analyse dévoile une réalité souvent ignorée : une partie significative des fonctionnaires chinois, appartenant à *l'International Poverty Reduction Center*, au ministère du commerce extérieur et au ministère des affaires étrangères, sont sensibles aux « bonnes pratiques » de l'efficacité de l'aide et souhaitent les diffuser dans leurs politiques nationales. Une représentante du CAD travaillant dans le secteur de l'efficacité de l'aide témoigne : « *le groupe d'étude sur la Chine du CAD-OCDE a organisé une journée de séminaire sur l'évaluation. La Chine envoie des stagiaires au CAD-OCDE pour travailler sur le thème de l'évaluation. Elle organise des séminaires sur la question* »[1075]. L'AFD a également mené des activités de renforcement des capacités en évaluation auprès de fonctionnaires chinois sous la forme de séminaires ou de travaux d'évaluation où la Chine était observatrice. Les évaluateurs de l'AFD estiment que ces expériences ont permis un échange constructif sur l'évaluation avec les représentants chinois impliqués. Ces derniers ont semblé être intéressés et avoir la volonté de développer la pratique de l'évaluation dans leurs organismes de coopération pour le développement[1076].

Au niveau international, les opérateurs chinois de l'aide au développement sont donc réceptifs aux « bonnes pratiques » des donateurs occidentaux en termes d'efficacité de l'aide et d'évaluation et semblent ouverts à l'apprentissage mutuel. De cette analyse, nous concluons que le refus proclamé par la Chine de s'engager avec les donateurs traditionnels dans des accords internationaux relève plus d'une volonté de revendiquer une autonomie et une identité propre en tant que donateur dans le régime international de l'aide, que d'un rejet complet des bonnes pratiques occidentales. Ainsi, si le régime international de l'aide ne peut plus ignorer les revendications des donateurs émergents, la Chine aussi est affectée par les bonnes pratiques

[1073] *Ibid*.
[1074] BRANT, Philippa, *No Strings Attached: Chinese Foreign Aid and its Implications for the International Aid Regime*, Submitted in total fulfilment of the requirements of the degree of Doctor of Philosophy, School of Social and Political Sciences, Faculty of Arts, The University of Melbourne, mai 2012, 273p, p. 196.
[1075] Entretien avec une évaluatrice du CAD-OCDE, 2 juillet 2014.
[1076] Entretiens informels avec des personnels de l'unité d'évaluation de l'AFD lors de notre observation participante en 2012 et 2013.

de ce régime. Un double jeu d'influence a donc lieu et nous pouvons supposer que les donateurs se dirigent vers une convergence des pratiques.

D'autre part, bien que la Chine condamne le déclassement qu'elle subit de la part des donateurs occidentaux, elle joue de ce statut d'*outsider* pour faire valoir une approche alternative de la coopération pour le développement auprès des pays du Sud. Sa mise à l'écart lui sert donc auprès d'un certain nombre de pays bénéficiaires de l'aide. La Chine peut se permettre cette stratégie d'acteur isolé en raison de son volume important d'aide au développement qui lui donne le pouvoir de faire entendre sa voix aux donateurs occidentaux pouvant se sentir menacés par sa concurrence normative sur le terrain de l'aide.

Le régime de l'aide au développement est actuellement en profonde mutation avec une influence mutuelle des donateurs émergents et des donateurs traditionnels qui construisent de nouveaux référentiels de l'efficacité du développement et de l'évaluation. L'évaluation doit être partenariale, porter sur le développement dans son ensemble et être utile aux pays bénéficiaires.

Toutefois, bien que les donateurs émergents non-membres du CAD-OCDE soient, dans une certaine mesure, influencés par les « bonnes pratiques » de l'agenda sur l'efficacité de l'aide et de la Déclaration de Busan, ils « vendent » leur aide aux pays émergents en argumentant qu'ils ont des pratiques opposées à celles des donateurs traditionnels. Leur valeur ajoutée résiderait dans le fait que leur aide est plus souple, moins technocratique, avec moins de procédures de suivi-évaluation et de conditionnalités.

Les donateurs émergents gardent donc aujourd'hui une distance avec les « bonnes pratiques » de l'efficacité de l'aide des donateurs traditionnels, notamment en matière d'évaluation. Alors que les donateurs, tels que la Corée du Sud, le Brésil ou le Mexique, cherchent à conjuguer leurs approches de l'aide avec celles des donateurs traditionnels, l'Inde et la Chine sont davantage sur la défensive. Notre ambition est de voir comment ces positionnements ambivalents se traduisent concrètement sur le terrain de l'aide.

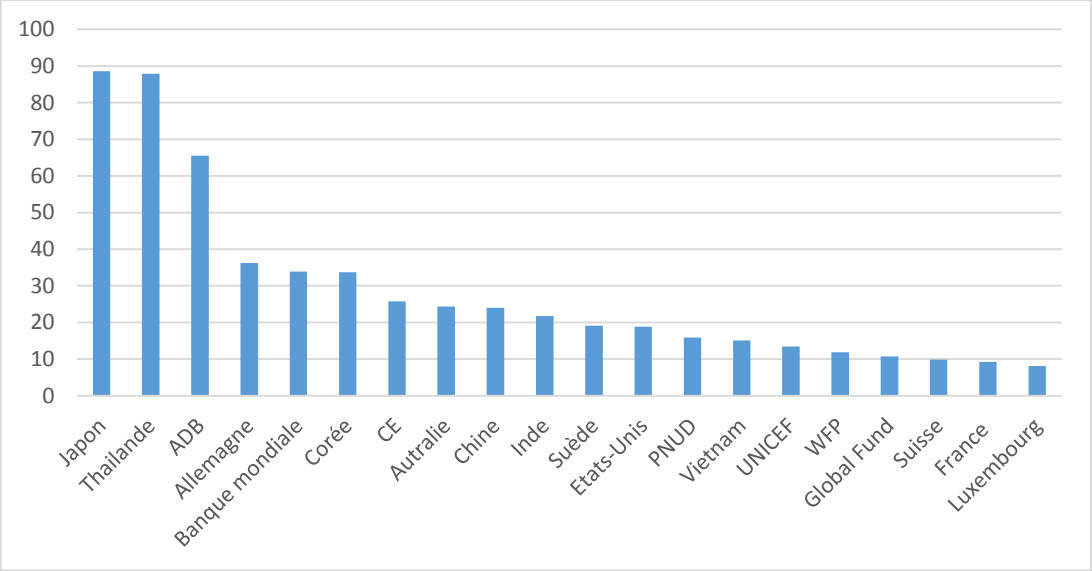
C. Des divisions politiques au niveau local

Notre point de départ pour cette analyse est notre expérience d'observation participante au sein de l'agence de l'AFD au Laos[1077], durant laquelle nous avons pu observer le jeu d'acteurs entre les donateurs traditionnels et émergents, notamment la Chine, la Corée du Sud et la Thaïlande. Le Laos est un terrain spécifique, dans le sens où les bailleurs émergents ont un poids très important qui concurrence en volume celui des donateurs traditionnels. Ils ont, en conséquence, une plus grande marge de manœuvre pour rester en dehors du régime de l'aide des bailleurs traditionnels et de ses bonnes pratiques. Nous confronterons ensuite les conclusions de notre étude sur le Laos à l'examen d'un terrain africain, le Sénégal, où les donateurs émergents sont moins implantés historiquement.

- **Le cas d'étude du Laos : une polarisation des acteurs autour des bonnes pratiques sur l'efficacité de l'aide**

Les donateurs émergents sont très présents au Laos, comme le montre le graphique ci-dessous :

Graphique 6 : Les 20 premiers donateurs du Laos (MUSD)



Source : LPDR Lao The National Roundtable Process, 2012

[1077] Cette phase d'observation participante a durée pendant deux mois. Nous étions chargées de préparer l'évaluation de la mise en œuvre des diligences environnementales et sociales d'un grand barrage sur un affluent du Mékong.

Parmi les 20 premiers donateurs du Laos, nous retrouvons la Thaïlande, la Corée du Sud, la Chine, l'Inde et le Viêtnam. Si ces donateurs ont toujours été présents au Laos, les volumes d'aide qu'ils apportent ont fortement augmenté durant les années 2000. L'aide de la Corée du Sud, par exemple, est passée de 3,64 millions de dollars en 2004 à 33,69 millions de dollars en 2009[1078], tandis que l'aide de l'Inde est passée de 1 million de dollars en 2004 à 21,74 millions de dollars en 2009[1079]. Les donateurs émergents ont donc un poids assez élevé pour concurrencer les modèles de développement diffusés par les donateurs traditionnels auprès des autorités nationales laotiennes.

Les donateurs émergents, bien que signataires de la Déclaration du Millénaire sur le Développement, interviennent peu sur les secteurs prioritaires des OMD. Ils investissent principalement dans l'infrastructure (construction de routes, de zones résidentielles, de centres de congrès, etc.), dans la production d'énergie (barrages hydroélectriques) et dans l'industrie, alors que les donateurs occidentaux délaissent ces secteurs d'activité pour se focaliser sur ceux des OMD (lutte contre la pauvreté, santé, éducation, etc.). Le gouvernement laotien vante cette « complémentarité » entre les donateurs[1080], les projets financés par les donateurs émergents étant plus rentables et permettant une croissance économique plus rapide que les actions des donateurs traditionnels sur l'égalité homme-femme, la santé maternelle ou la bonne gouvernance.

En 2006, le gouvernement du Laos a élaboré la Déclaration de Vientiane sur l'efficacité de l'aide[1081] qui reprend les principes de la Déclaration de Paris, en les « adaptant au contexte national du Laos »[1082]. Cette mesure a permis au Laos de se classer parmi les pays bénéficiaires désignés comme des « bons élèves » qui s'engagent à respecter la Déclaration de Paris. Au-delà de cette volonté du gouvernement de se « reclasser » parmi les « bons » bénéficiaires de l'aide, l'adoption de la Déclaration de Vientiane est le fruit d'une pression des donateurs qui voulaient voir les officiels locaux mettre en place un cadre de dialogue et surtout de contrôle sur la gestion et les résultats de l'aide au développement.

La Déclaration de Vientiane a été signée par la Chine, la Corée du Sud et la Thaïlande ; l'Inde et le Viêtnam ayant préféré s'abstenir. La Chine, qui n'a pas adhéré à la Déclaration de Paris au niveau international, l'a donc ratifiée au niveau local. Pour justifier cette ambivalence,

[1078] www.rtm.org.la, consulté le 27 mars 2014.
[1079] www.rtm.org.la, consulté le 27 mars 2014.
[1080] www.rtm.org.la, consulté le 14 avril 2014.
[1081] Government of the Lao People's Democratic Republic and the Partners in Development, *Vientiane Declaration on Aid Effectiveness*, 2007.
[1082] www.rtm.org.la; consulté le 27 mars 2014.

les officiels chinois expliquent qu'ils ne sont pas contre la Déclaration de Paris en elle-même, mais contre le fait que les donateurs imposent un ensemble de conditionnalités aux pays bénéficiaires. Dans le cas du Laos, où le gouvernement a élaboré la Déclaration de Vientiane, en collaboration avec les donateurs occidentaux et où il est demandeur d'un engagement de la Chine à la respecter, les représentants chinois ont répondu favorablement[1083]. Selon nous, la principale raison est plutôt que la ratification de cette déclaration locale entraîne beaucoup moins d'engagements et de contraintes que la signature d'un accord international rendant la Chine redevable de ses agissements en matière d'aide au développement envers l'ensemble des acteurs du développement.

De plus, la ratification par la Chine de la Déclaration de Vientiane ne signifie pas qu'elle est pleinement engagée dans les exercices de coordination et d'harmonisation entre les donateurs, notamment en matière de suivi-évaluation. Nous retrouvons là la situation de lu partenariat global pour lequel la Chine n'a pas adhéré au cadre de suivi-évaluation. En 2005, le Laos a mis en place dix groupes de travail sectoriels pour permettre aux donateurs et aux ministères d'échanger sur les réalisations et les résultats de l'aide, afin d'augmenter l'efficacité de l'aide et de mieux répondre aux besoins du gouvernement[1084]. Or, les donateurs émergents ne participent pas aux réunions de ces groupes sectoriels. Un officiel japonais déclare, par exemple, « *dans les années récentes, la Chine a rapidement augmenté son aide économique au Laos et gagné du poids en tant que donateur. Toutefois, elle ne participe pas aux réunions des donateurs et la nature de son aide est loin d'être claire* »[1085]. Selon un représentant de l'AFD au Laos, « *les Chinois sont complètement absents des structures de coordination et d'harmonisation de l'aide. Ces structures sont des émanations du CAD et seuls les donateurs membres du CAD y participent. Il n'y a aucun dialogue avec les donateurs chinois, nous ne les voyons même pas* »[1086]. La Corée du Sud fait figure d'exception parmi les donateurs émergents et participe à ces groupes sectoriels depuis 2010. Ses pratiques sur le terrain ont donc évolué en vue et à la suite de son intégration au CAD-OCDE, tout du moins en ce qui concerne la transparence de son aide, la coordination et l'harmonisation avec les autres donateurs. La Corée du Sud évalue désormais les programmes d'aide qu'elle met en place au Laos en essayant de se

[1083] Entretien avec un développeur chinois au Laos, 18 juillet 2012.
[1084] www.rtm.org.la, consulté le 27 mars 2014.
[1085] Entretien avec un membre du bureau de la JICA au Laos, 25 juin 2012.
[1086] Entretien avec un responsable de projets de l'AFD au Laos, 30 juin 2012.

conformer aux standards d'évaluation du CAD-OCDE[1087]. Les résultats de ces évaluations sont partagés avec les autres donateurs lors des réunions de coordination sectorielles.

Les autres donateurs émergents participent seulement aux tables rondes annuelles qui réunissent l'ensemble des donateurs du Laos. Ils fournissent alors des informations sur leur aide au développement, bien que, dans le cas de la Chine et de la Thaïlande, elles soient parcellaires, ne couvrant pas toutes les activités de l'aide et n'étant pas détaillées au niveau des programmes et des projets[1088]. De plus, la Chine et la Thaïlande ont refusé d'intégrer « la plateforme du suivi de l'aide », créée par le gouvernement du Laos en 2011, qui a pour objectif de i) fournir un compte-rendu détaillé de l'aide publique au développement au Laos incluant la coopération Sud-Sud ; ii) souligner l'impact des efforts du gouvernement pour soutenir la mise en œuvre de l'agenda sur l'efficacité de l'aide sur tout le territoire du Laos ; iii) et fournir un vecteur pour une discussion efficace sur les stratégies d'aide des donateurs. Cet engagement à réaliser un suivi régulier et transparent de leur aide au développement allait trop loin pour les donateurs chinois et thaïlandais.

L'adhésion des donateurs émergents à la Déclaration de Vientiane sur l'efficacité de l'aide relève donc surtout du symbole et de la rhétorique. Elle s'illustre par leur participation aux « grands messes » annuelles réunissant la « communauté de l'aide » du Laos. À côté, les donateurs émergents conservent des procédures très faibles de suivi-évaluation, de transparence et de redevabilité[1089]. Le fait que le gouvernement du Laos ait élaboré une version locale de la Déclaration de Paris, ratifiée par les donateurs émergents, aurait pu inciter ces derniers à la respecter *a minima*. Toutefois, les échos que nous avons pu avoir quant à cette hypothèse sont assez sceptiques : « *le Laos est signataire de la Déclaration de Paris et a institutionnalisé la Déclaration de Vientiane sur l'efficacité de l'aide. C'est un pas en avant vers une gestion de l'aide plus harmonisée et plus transparente, dans laquelle les projets sont suivis et évalués selon des indicateurs et des objectifs cohérents, d'une part entre les donateurs et, d'autre part avec les priorités du gouvernement. Toutefois, les problèmes de corruption au Laos font que la Déclaration n'est pas appliquée par les officiels locaux en charge des projets d'aide. De plus, cet engagement du Laos n'a eu aucun effet sur les pratiques des donateurs émergents, car le pays ne jouit pas d'une puissance de négociation et de contradiction suffisante envers les*

[1087] Certains des rapports d'évaluation sont disponibles sur les sites internet de la KOICA : http://www.koica.go.kr/english/schemes/evaluation/report/index.html et de l'EDCF : http://www.edcfkorea.go.kr/edcfeng/bbs/pubanew/list.jsp?bbs_code_id=1368577018458&bbs_code_tp=BBS_4, consultés le 5 décembre 2014.

[1088] Observations faites durant notre observation participante au bureau de l'Agence française de développement au Laos de juin à août 2012.

[1089] Observations faites durant notre observation participante au bureau de l'AFD au Laos de juin à août 2012.

grandes puissances asiatiques voisines, comme la Chine, la Thaïlande et le Viêtnam »[1090]. En effet, le Laos entretient des liens diplomatiques et de dépendance forts avec ces pays qui financent une large part de son développement économique. L'effet d'entraînement créé par l'institutionnalisation de la Déclaration de Vientiane sur les pratiques des donateurs émergents et les officiels du Laos est donc faible.

À travers le cas d'un projet de grand barrage hydroélectrique, nous allons désormais étudier les effets de reclassement que produit indirectement cette diversité des pratiques entre les donateurs occidentaux et les donateurs émergents.

Selon les bonnes pratiques occidentales, la construction de grands barrages est précédée d'évaluations *ex ante*[1091] qui peuvent durer une dizaine d'années, d'autant plus lorsque les projets impliquent des déplacements de population indigène. Les bailleurs de fonds mettent en place des programmes de compensation très sophistiqués pour dédommager ces populations déplacées, allant de la reconstruction des villages, à l'apport d'infrastructures (routes, irrigation, etc.) supplémentaires, jusqu'à des activités de renforcement des capacités et d'amélioration de la santé publique. Ces programmes de compensations font l'objet d'un suivi-évaluation très rigoureux. Par exemple, le cadre de suivi-évaluation des compensations sociales du barrage au Laos sur lequel nous avons travaillé lors de notre observation participante à l'AFD comptait plus de 150 indicateurs, destinés à suivre et à évaluer la bonne mise en œuvre et les effets du déplacement et de la réinstallation des 6 500 personnes déplacées. Ces indicateurs allaient jusqu'à énoncer le nombre de pelles et de couteaux à fournir aux villageois[1092]. Des actions de limitation des effets environnementaux du barrage avaient également été mises en place, assorties elles aussi d'un nombre important d'indicateurs.

Le dispositif de suivi-évaluation a donc été lourd à mettre en œuvre, aussi bien pour l'opérateur du projet que pour les bailleurs de fonds et le gouvernement. Environ 170 missions de suivi ont été organisées en sept ans pour contrôler les progrès effectués vers l'atteinte des objectifs environnementaux et sociaux du projet de barrage. Ces objectifs visaient notamment à ce que tous les foyers déplacés atteignent le seuil de pauvreté national[1093]. Or, les populations déplacées avaient pour la plupart, à l'origine, un niveau de vie largement inférieur à ce seuil. Ainsi, l'ingénierie économique, sociale et environnementale mise en œuvre pour augmenter les

[1090] Entretien avec un chef de projets de l'agence de l'AFD au Laos, 30 juin 2012.
[1091] Pour une définition de ce terme, voir lexique en annexe 7.
[1092] Document confidentiel
[1093] Document confidentiel.

capacités et les revenus des populations déplacées a mobilisé plus d'énergie que le projet de barrage en lui-même.

Nous avons observé que la lourdeur du mécanisme de suivi-évaluation des effets environnementaux et sociaux du barrage a entraîné des tensions entre les bailleurs de fonds et le gouvernement laotien. Le gouvernement trouvait ces exigences trop coûteuses et incohérentes avec le contexte local. Il arguait que le respect des objectifs déterminés par les bailleurs de fonds risquait de créer une situation de décalage économique et social entre les bénéficiaires des compensations et les autres populations de la région, ce qui favoriserait l'émigration vers cette zone et des demandes de mise à niveau des conditions de vie dans les zones avoisinantes, voire dans l'ensemble du pays[1094]. De plus, les délais et les coûts supplémentaires entraînés par les évaluations *ex ante, in itinire* et *ex post* ont suscité le mécontentement du gouvernement du Laos, dont la priorité était la construction rapide du barrage.

Malgré ces tensions, ce projet a été l'opportunité pour le gouvernement du Laos de définir ses premiers principes normatifs en matière de gestion et de suivi-évaluation des déplacements de population. En 2005, le Premier ministre du Laos a publié un décret sur « les compensations et la réinstallation des personnes affectées par des projets de développement »[1095] qui présente les mesures à suivre en termes d'évaluation *ex ante* des risques économiques et sociaux dans le cas de déplacement de population. En 2006, le gouvernement a élaboré un document de politique nationale sur « la durabilité économique et sociale du secteur hydrologique au Laos »[1096]. Ce document présente des bonnes pratiques à suivre en matière de suivi-évaluation environnemental, économique et social, de consultations avec les populations locales, de transparence, de respect des normes internationales et de redevabilité. Incité par les donateurs, le gouvernement du Laos commence donc à élaborer des documents normatifs sur la gestion des risques, la protection de l'environnement, l'évaluation et la transparence.

Toutefois, les capacités en évaluation des officiels locaux sont faibles. Dans le cadre des projets de grandes infrastructures, les bailleurs de fonds et leurs opérateurs techniques mettent parfois en place des mécanismes de renforcement des capacités en suivi-évaluation. Par exemple, dans le cas de notre projet de grand barrage, une équipe mixte, composée d'experts internationaux et laotiens, était chargée de suivre et d'évaluer les effets environnementaux du

[1094] Propos recueilli auprès d'opérationnels ayant travaillé sur le projet du barrage lors de notre mission de terrain. En raison de la sensibilité du projet, les noms des interviewés ne peuvent être rendus publics.
[1095] *Ibid.*
[1096] WORLD BANK, 2010, *op. cit.*

barrage. La création de ce système mixte de suivi-évaluation a été facilitée par l'existence d'une institution nationale spécialement dédiée aux enjeux environnementaux, la *Science, Technology and Environment Organisation,* dont le personnel a pu être mobilisé. Pour le suivi-évaluation des effets économiques et sociaux, le renforcement des capacités a été plus compliqué, car il n'existe pas, au Laos, d'institution nationale chargée spécifiquement de ces enjeux. Pour dépasser ce problème, le gouvernement a créé des structures *ad hoc* chargées du suivi et de l'évaluation du processus de réinstallation des populations déplacées. Mais, étant donné qu'il n'y a pas d'organisation nationale pour se substituer à ces structures à la fin du projet, le renforcement des capacités du personnel risque de disparaître une fois les donateurs partis[1097]. Ce risque de perte des compétences transmises est d'autant plus grand que le gouvernement préfère désormais se tourner vers des financeurs moins exigeants, comme la Chine ou la Thaïlande, pour la construction de grands barrages. En témoigne le fait que, depuis 2005, sur la dizaine de nouveaux barrages construits au Laos, les financements proposés par les bailleurs de fonds traditionnels ont quasiment tous été rejetés en faveur de ceux des donateurs émergents[1098].

Ce détournement du gouvernement laotien vers les donateurs émergents amène les donateurs occidentaux à penser que « *dans le secteur des barrages, ils sont allés trop loin en termes de procédures environnementales et sociales* »[1099]. Selon un représentant de l'AFD, « *cette concurrence des bailleurs émergents fait revoir à la baisse les exigences des donateurs traditionnels en matière de limitation des effets environnementaux et sociaux des projets de grandes infrastructures* »[1100]. Leur objectif est désormais de trouver un juste équilibre entre l'ingénierie environnementale et sociale, l'efficacité et l'efficience économique des projets de grands barrages. Sur le terrain du Laos, par peur d'un déclassement, les donateurs traditionnels tendent donc à revoir à la baisse leurs exigences de conditionnalités, de contrôle et de suivi-évaluation. La montée en puissance des donateurs émergents, qui ont des référentiels différents de celui des donateurs traditionnels, a donc bien un effet significatif sur les pratiques de ces derniers.

En effet, les donateurs émergents, notamment la Chine et la Thaïlande, posent moins de conditionnalités que les donateurs traditionnels à la construction de barrages hydroélectriques. Les populations sont déplacées avec le minimum de compensations et les évaluations *ex ante*

[1097] Entretien mené avec un membre de l'entreprise chargée de la construction du barrage et de la mise en œuvre des compensations environnementales et sociales lors de notre mission de terrain au Laos en juin 2012.
[1098] HAFTERDORN, Esther, *État des lieux du secteur de l'hydroélectricité au Laos et identification des opportunités de financement pour l'AFD,* AFD, Bureau de Vientiane, RDP Lao, novembre 2013.
[1099] Réunion de travail sur l'évaluation des impacts environnementaux et sociaux du barrage X, à Vientiane, le 5 juin 2012.
[1100] *Ibid.*

des conséquences environnementales et sociales des barrages sont plus une formalité qu'un réel enjeu pouvant annuler les projets[1101]. Sur le terrain, nous constatons que les principes d'équité, de bénéfice mutuel et d'appropriation, prônés par les donateurs émergents, s'adressent aux gouvernements partenaires et non à leur population. La priorité du gouvernement laotien est le regroupement des populations indigènes pour lutter, officiellement, contre la production et le trafic de drogue et, officieusement, contre le désordre politique et social[1102]. Les donateurs émergents suivent cette ligne de conduite lorsque des projets de développement entraînent un déplacement de population, même si elle entraîne des conséquences très négatives sur le développement économique et social des personnes déplacées. Aussi bien les donateurs émergents que le Laos trouvent leur compte dans cette gestion moins contraignante des aspects environnementaux et sociaux des barrages. Le Laos, d'une part économise des négociations, des réformes de sa gouvernance environnementale et sociale, du temps et de l'argent, les donateurs émergents, d'autre part construisent des infrastructures à moindres coûts, plus rapidement et peuvent ensuite bénéficier de l'électricité produite selon leur approche de « bénéfices mutuels ».

Les donateurs traditionnels et émergents travaillent donc chacun de leur côté sur le terrain du Laos sans échanger de « bonnes pratiques » de suivi ou d'évaluation. À ce jour, il n'y a eu aucune expérience de renforcement des capacités en évaluation ou de coopération « triangulaire » entre donateurs du Nord et du Sud. Les donateurs occidentaux que nous avons interrogés estiment ce type de coopération peu probable malgré leur ouverture à cette possibilité. Ils pensent que les donateurs émergents ne veulent pas vraiment travailler conjointement avec les autres donateurs, qu'ils fuient ce type d'opportunités. Ils n'envisagent pas de possibles coopérations « *avec des pays dont les pratiques sont aussi peu transparentes et dont la culture est non-démocratique* ». Les différences, jugées insurmontables, dans la nature et la gestion de l'aide sont également soulignées[1103]. Le sentiment d'un éloignement culturel fort entre les donateurs occidentaux et émergents est donc très présent, à tel point qu'il est vu comme un blocage à la mise en œuvre d'un travail conjoint. Il est renforcé par le fait que les bailleurs émergents prônent de plus en plus une « culture asiatique » du développement. Ce sentiment est toutefois à relativiser, dans la mesure où il a suffi à la Corée du Sud, jusqu'alors considérée comme ayant des pratiques antinomiques à celles des donateurs traditionnels,

[1101] Ces constats sont tirés de l'analyse des conditions de mise en place des projets de barrages financés par les donateurs émergents depuis 2005 qui seraient trop longue à décrire dans cette recherche.
[1102] BAIRD, Ian, G, SHOEMAKER, Bruce, "Unsettling Experiences: Internal Resettlement and International Aid Agencies in Laos", *Development and Change,* vol. 38, n°5, 2007, pp. 865-888.
[1103] Entretien avec un représentant d'une agence d'aide occidentale au Laos, 3 juillet 2012.

d'afficher une volonté de réforme et de rentrer au CAD-OCDE pour qu'elle soit considérée comme un « pair » par ces derniers. Les bonnes pratiques, leur respect ou leur rejet, sont donc un marqueur d'identification normative fort pour les donateurs.

Sous couvert d'une adhésion formelle à la Déclaration de Vientiane sur l'efficacité de l'aide et d'un effort de transparence *a minima*, la Chine développe une approche de l'efficacité de l'aide divergeant fortement de celle des donateurs occidentaux. Ils évaluent peu leurs programmes d'aide au développement, se limitant à une mesure quantitative des réalisations effectuées et des fonds alloués[1104]. Cette situation convient au gouvernement du Laos qui, malgré l'élaboration de la Déclaration de Vientiane et ses efforts pour en respecter les standards, voit en l'aide des bailleurs émergents une manne de financements faciles et ayant un moindre coût politique. Les représentants des rares ONG laotiennes, quant à eux, semblent préférer les méthodes des donateurs occidentaux qui leur offrent plus de sécurité, notamment en matière de risques environnementaux et sociaux[1105]. Toutefois, dans un pays comme le Laos, où la liberté d'expression est faible, la société civile n'a pas le pouvoir de critiquer ouvertement les bailleurs émergents. Sur le terrain, il apparaît que ces derniers n'ont, aujourd'hui, pas d'intérêts à augmenter leurs pratiques d'évaluation et à être plus transparents.

Au Laos, le fait que les donateurs traditionnels augmentent leurs bonnes pratiques de redevabilité et d'évaluations, alors que les donateurs émergents n'empruntent pas cette voie, entraîne le déclassement relatif des donateurs traditionnels par rapport aux donateurs émergents. Toutefois, cette situation pourrait être due au fait que le Laos est un pays où l'influence des donateurs émergents asiatiques est forte. Ainsi, nous allons maintenant analyser si cette situation s'observe également sur un terrain africain.

- **Le cas d'étude du Sénégal : une remise en cause du reclassement des donateurs émergents ?**

Afin d'élever notre réflexion au-delà du seul cas du Laos, nous avons voulu aborder un autre cas d'étude, dans un contexte très différent, le Sénégal, où les donateurs émergents semblent avoir moins de poids qu'au Laos[1106].

[1104] Entretien avec un chef de projets de l'AFD au Laos, 30 juin 2012.
[1105] BAIRD, Ian. G, SHOEMAKER, Bruce 2007, *art. cit.*
[1106] Dans ce cas d'étude, nous nous concentrerons sur le cas de la Chine vu que la Corée du Sud ne met pas en œuvre une action significative d'aide au développement au Sénégal. Cela ne nous est pas apparu gênant dans la mesure où nous avons démontré dans le cas du Laos que la Corée du Sud avait renforcé ses pratiques d'évaluation, sur le terrain, après son adhésion au CAD-OCDE.

La montée en puissance des donateurs émergents au Sénégal est relativement récente et la Chine apparaît, de loin, le donateur le plus dynamique et le plus visible d'entre eux. Les relations entre la Chine et le Sénégal ont débuté en 2005, lorsque le président Abdoulaye Wade s'est tourné vers la Chine au détriment de Taïwan. En 2013, la Chine a accordé plus de 170 millions d'euros d'aide au Sénégal, ce qui en fait le quatrième bailleur de fonds du pays après la Banque mondiale, la Banque africaine de développement et la France[1107].

Comme au Laos, les donateurs émergents ne participent pas aux structures de coordination entre les bailleurs de fonds. Un représentant d'une agence d'aide européenne témoigne à ce sujet : « *les seules informations que l'on a sur l'aide des donateurs émergents, c'est par la presse, ou un peu par hasard sur le terrain. On ne les voit pas du tout, et c'est une attitude générale dans les pays africains où j'ai pu travailler. Trouver des chiffres sur l'aide chinoise est difficile. Nous n'avons pas d'acharnement sur la question, mais parfois nous aimerions savoir et nous ne trouvons pas grand-chose* »[1108]. Ce non-alignement se concrétise de manière encore plus visible qu'au Laos, dans la mesure où le gouvernement du Sénégal a créé des services spéciaux au sein de ses ministères pour traiter l'aide des donateurs émergents[1109]. Il existe donc, concrètement, un double système de gestion de l'aide au Sénégal : celui des donateurs du CAD-OCDE et des organisations multilatérales, et celui des donateurs émergents[1110]. Cette situation suscite l'inquiétude des donateurs occidentaux qui considèrent qu'elle sape leurs efforts d'harmonisation et de coordination de l'aide.

Nous n'avons pas connaissance d'évaluations qui auraient été mises en œuvre par les donateurs émergents au Sénégal. Selon un représentant de la société civile sénégalaise travaillant avec des donateurs émergents, « *les questions de l'évaluation ne sont pas au cœur de la coopération entre les donateurs émergents et le Sénégal. Ce n'est pas une question qui les intéresse. Je n'ai jamais entendu parler d'activités des donateurs émergents ayant fait l'objet d'évaluations. Ces donateurs interviennent pour retirer de la valeur ajoutée de leurs projets. Ils ont une logique de marché, l'évaluation n'est pas leur problème* »[1111]. Les seules « évaluations » qui seraient mises en œuvre relèveraient davantage de l'étude de marché ou du calcul de la rentabilité des projets, plutôt que de l'analyse de leurs résultats sur le

[1107] AUREGAN, Xavier : http://www.rfi.fr/europe/20140219-chine-senegal-macky-sall-xi-jinping-visite/, consulté le 28 avril 2014.
[1108] Entretien avec le directeur de l'AFD au Sénégal, 4 avril 2014.
[1109] DAC-OECD, *Évaluation de la Déclaration de Paris, Évaluation pays du Sénégal*, 2011.
[1110] CAD-OCDE, *Évaluation de la Déclaration de Paris, Cas d'étude du Sénégal*, 2011.
[1111] Entretien avec un consultant sénégalais en évaluation, responsable de la société sénégalaise d'évaluation, 4 avril 2014.

développement national. Au niveau local, nous n'avons pas relevé d'échanges entre les donateurs émergents et traditionnels en matière d'évaluation.

Les personnes que nous avons interrogées, aussi bien du côté des donateurs, que du côté des officiels sénégalais, observent que la montée en puissance des donateurs émergents tend à remettre en cause les principes de l'efficacité de l'aide. Les rédacteurs de l'évaluation de la Déclaration de Paris au Sénégal notent que « la lenteur des réformes et des changements (issus de la Déclaration de Paris), tant au niveau du gouvernement que des partenaires de développement, pose un certain nombre de risques, dont l'essoufflement et le désintérêt des acteurs envers la poursuite de la mise en œuvre de la Déclaration de Paris. Cela surtout en regard de la montée en puissance des donateurs non-signataires de la Déclaration, comme la Chine ou l'Iran »[1112]. Ce désintérêt de la Déclaration de Paris se retrouve dans le discours des acteurs occidentaux. Ainsi, un représentant de l'AFD déclare : « *sur ces histoires de Déclaration de Paris et compagnie, nous avons le sentiment, sur le terrain, que ces trucs-là sont très technocratiques. Ce sont des émanations de la bureaucratie internationale qui se réunit dans des grands forums parce qu'elle a soudainement pris conscience que l'aide devait gagner en efficacité. Ces accords n'ont plus vraiment de sens dans l'écosystème actuel de l'aide, de plus en plus foisonnant, avec des acteurs multiples qui ne se coordonnent pas. De plus, nous avons le sentiment que les pays bénéficiaires s'en fichent un peu, qu'ils sont peu demandeurs de ces cadres de suivi-évaluation de l'aide* »[1113].

Au Sénégal également, le gouvernement se détourne des donateurs traditionnels au profit des donateurs émergents. Les justifications souvent données à ce choix sont que *« le financement des pays non traditionnels est plus souple, plus rapide à obtenir et moins contraignant en termes de bonne gouvernance et de suivi-évaluation que celui des pays du Nord* »[1114]. Un autre avantage cité est que l'aide des donateurs émergents ne s'inscrit pas dans un cadre de domination du Nord sur le Sud : « *le cadre de coopération entre la Chine et le Sénégal est différent. Avec les donateurs classiques, nous avons une relation inscrite dans une histoire très complexe qui a déteint sur les coopérations bilatérales. La Chine ne s'inscrit pas dans ce sillage historique où le Nord développé apprend au Sud sous-développé comment gouverner. Son rapport avec le Sénégal est beaucoup plus pragmatique* »[1115]. Le ministère des infrastructures déclare que la perspective de coopérer avec les donateurs émergents est

[1112] DAC-OECD, *Évaluation de la Déclaration de Paris, cas d'étude du Sénégal,* 2011.
[1113] Entretien avec le directeur de l'agence de l'AFD au Sénégal, 3 avril 2014.
[1114] Entretien avec un agent du ministère de l'environnement et du développement durable du Sénégal, 5 avril 2014.
[1115] *Ibid.*

alléchante, car l'aide des donateurs traditionnels est lourde à gérer et implique plein de « complications »[1116]. Malgré les progrès en termes d'alignement et d'appropriation, le ministère souffrirait toujours des coûts de gestion élevés de l'aide classique. La hausse des exigences en termes de suivi et d'évaluation des programmes d'aide, à des fins de transparence et de redevabilité, est perçue comme contribuant, en grande partie, à cette lourdeur bureaucratique de l'aide des donateurs du CAD-OCDE[1117].

Comme c'est le cas au Laos, le gouvernement sénégalais bénéficie de l'écart normatif entre les donateurs traditionnels et émergents. Un interviewé sénégalais relate : « *les donateurs émergents viennent, ils construisent des bâtiments, ils se positionnent surtout sur des questions d'infrastructures. Ils ont une approche pragmatique du développement, ils apportent des solutions matérielles à des problèmes matériels. Pendant ce temps, nos partenaires classiques sont sur des conditionnalités et des enjeux politiques comme l'amélioration de la gouvernance, la mise en place de l'État de droit, le renforcement des capacités de la société civile, le renforcement de la démocratie. Pour moi, ce sont deux types de coopération complémentaires* »[1118]. Ainsi, les donateurs émergents financent des projets de prestige (une vingtaine de stades dans le pays, le Palais national, le Musée des arts africains, etc.), d'infrastructures (notamment des routes) et de génie civil (particulièrement des barrages). L'aide chinoise est particulièrement visible, en raison de la présence d'entreprises et d'une main d'œuvre nombreuse sur les chantiers d'aide, ainsi que de la mise en place de plaques commémoratives, de l'affichage du drapeau chinois, et de grandes déclarations dans la presse. Les donateurs occidentaux, quant à eux, se concentrent sur les domaines prioritaires des OMD.

Les perspectives différenciées des bailleurs permettent au gouvernement sénégalais d'avoir un levier de négociation avec les bailleurs de fonds occidentaux pour qu'ils soient « plus souples ». C'est ce que relève un donateur : « *le gouvernement sénégalais veut créer une certaine compétition entre les donateurs et recueillir le maximum de financements d'où qu'ils viennent. Au Sénégal, comme dans d'autres pays, le gouvernement joue une stratégie pour mettre en concurrence les bailleurs de fonds, surtout quand il s'agit des bailleurs de fonds émergents contre les bailleurs de fonds traditionnels* »[1119]. Ainsi, contrairement à notre hypothèse de départ, les bailleurs occidentaux déclarent, au Sénégal également, ressentir un effet de déclassement à cause de la concurrence des donateurs émergents.

[1116] CAD-OCDE, *Évaluation de la Déclaration de Paris, cas d'étude du Sénégal*, 2011.
[1117] DAC-OECD, *Évaluation de la Déclaration de Paris, cas d'étude du Sénégal,* 2011.
[1118] Entretien avec un évaluateur sénégalais, président de la société sénégalaise d'évaluation, 3 avril 2014.
[1119] Entretien avec le directeur de l'agence de l'AFD au Sénégal, 4 avril 2014.

Les donateurs traditionnels sont peu habitués à une telle situation de concurrence et la vivent péniblement[1120]. Un représentant d'une agence d'aide européenne témoigne : « *les donateurs chinois font des annonces pharaoniques (de projets d'aide) dans la presse que l'on découvre avec surprise, c'est assez déstabilisant surtout lorsqu'il s'agit de projets sur lesquels nous travaillons déjà depuis plusieurs années. Je peux citer deux exemples : le chemin de fer Dakar-Bamako qui est un dossier important sur lequel nous travaillons avec la Banque mondiale. Également, la grande interconnexion électrique du Sud-Est du pays sur laquelle nous travaillons avec la Banque européenne d'investissement et d'autres donateurs depuis quelques mois. Or, la Chine a annoncé qu'elle financerait ces deux projets sans que nous en ayons été préalablement informés* »[1121].

Si nous reprenons l'exemple du secteur des barrages hydroélectriques, les résultats de cette concurrence sont, aujourd'hui, flagrants au Sénégal. Les donateurs occidentaux ont complètement été évincés du financement de ce type de projets en faveur des donateurs émergents. De même qu'au Laos, c'est la souplesse des donateurs émergents en matière d'évaluation *ex ante* et *ex post*, de suivi et de conditionnalités qui a entraîné le gouvernement sénégalais à se détourner quelque peu des donateurs traditionnels.

Au Sénégal, nous n'observons pas de processus d'échanges de bonnes pratiques ou d'apprentissage mutuel en matière d'évaluation entre les donateurs traditionnels et les donateurs émergents. Un représentant d'une agence d'aide européenne décrit les relations qu'il entretient avec les donateurs chinois : « *l'influence normative en matière de suivi-évaluation entre les donateurs émergents et les donateurs traditionnels est très faible, car les contacts sont peu nombreux. On se fait de grands sourires, mais il n'y a pas vraiment de dialogue technique et pas de transparence sur ce que font les uns et les autres. Apparemment, les Chinois ne sont pas intéressés par ce genre d'échanges* »[1122]. Selon certains donateurs occidentaux, le risque est de voir l'efficacité de l'aide se réduire à la conception qu'en ont des donateurs émergents, c'est-à-dire à une efficacité de décaissement ou de moyens. Ils considèrent que cela reviendrait à un recul de plusieurs décennies dans la pratique de l'évaluation qui serait limitée à une analyse quantitative de l'aide allouée et des réalisations des projets d'aide. Un tel retour en arrière nous paraît peu probable.

[1120] EASTERLY, William, *The cartel of good intentions: the problem of bureaucracy in foreign aid, Center for Global Development*, octobre 2002.
[1121] Entretien avec un représentant de la Commission européenne au Sénégal, 7 avril 2014.
[1122] Entretien avec le directeur de l'agence de l'AFD au Laos, 3 avril 2014.

En effet, contrairement au cas du Laos, les acteurs occidentaux ne déclarent pas revoir à la baisse leurs exigences en termes de conditionnalités, de redevabilité et de suivi-évaluation en réaction à la concurrence des pays émergents : « *nous n'avons absolument pas de stratégie de baisse de nos exigences et de nos conditionnalités envers le gouvernement sénégalais face à la présence des donateurs émergents. Au contraire, il peut nous arriver, si on en a connaissance à temps, de nous mettre en travers de, ou de faire des remarques sur, tel ou tel projet qui serait financé par les donateurs émergents et qui ne correspondrait pas aux politiques prioritaires du pays, ni à des choix économiques ou écologiques appropriés* »[1123]. Ce témoignage laisse penser que les donateurs émergents ont tout de même moins de poids pour influer sur les bonnes pratiques des donateurs traditionnels dans des contextes où l'implantation historique, économique et diplomatique de ces derniers est plus solide.

Nous estimons également que la tendance générale est davantage à une hausse de la pratique de l'évaluation chez les donateurs émergents, même si elle ne correspond pas aux standards du CAD-OCDE, plutôt qu'à un recul des exigences de redevabilité et de transparence chez les donateurs occidentaux, comme en témoignent les échanges de pratiques au niveau international.

Nous défendons la thèse que la réticence des donateurs émergents à renforcer leur pratique de l'évaluation n'est pas la résultante de faiblesses techniques, mais bien d'un positionnement stratégique. Les donateurs chinois seraient tout à fait capables de développer des techniques de suivi-évaluation, d'améliorer la transparence et la responsabilité environnementale et sociale de leurs projets. Ils n'en ont tout simplement pas l'intérêt à ce jour, tant que les gouvernements seront attirés par leurs financements souples et peu conditionnés. Modifier leurs pratiques pour les faire correspondre aux standards des donateurs du CAD-OCDE et multilatéraux leur ôterait leur avantage comparatif. Au niveau national, la pression exercée par les donateurs occidentaux pour que les donateurs émergents rejoignent les réunions de coordination inter-donateurs et les cadres d'harmonisation de l'aide semble faible. Toutefois, nous pouvons penser que sur le long terme, les conséquences négatives des projets d'aide des donateurs émergents pourront amener les gouvernements à augmenter leurs exigences de transparence, de suivi et d'évaluation auprès des donateurs émergents, mais cela reste purement hypothétique aujourd'hui.

[1123] *Ibid.*

- **Conclusion de chapitre**

À travers le cas des donateurs émergents, ce chapitre a montré les jeux de pouvoir qui se nouent autour des « bonnes pratiques » de l'efficacité de l'aide, et plus spécifiquement de l'évaluation. L'adoption de pratiques normées d'évaluation, selon les principes édictés par le CAD-OCDE, est devenue un critère de qualité de l'aide. Au niveau international, les donateurs qui suivent ces principes sont considérés comme des « bons » donateurs, alors que les autres sont pointés du doigt comme étant des « mauvais » donateurs et s'estiment déclassés. À travers l'évaluation, et plus largement l'efficacité de l'aide, se confrontent des conceptions politiques différentes de l'aide au développement.

Ce phénomène de déclassement s'inverse toutefois au niveau local. En effet, sur le terrain de l'aide, ce sont davantage les donateurs occidentaux qui se retrouvent déclassés par rapport aux donateurs émergents, en raison du haut niveau de bureaucratisation, de conditionnalité et de technicité de leur aide, considérée comme trop contraignante par les pays bénéficiaires.

Par ailleurs, les positionnements des donateurs sur la scène internationale, s'affichant « pour » ou « contre » les « bonnes pratiques » d'évaluation, de redevabilité et de transparence, sont souvent caricaturaux. En effet, même sur la question de l'évaluation, qui est éloignée de la culture organisationnelle des organismes d'aide des pays émergents, les échanges se multiplient ces dernières années au niveau international, ce qui tend à montrer une volonté d'apprentissage de la part d'un nombre significatif de représentants des pays émergents.

Les donateurs émergents, même la Chine, commencent à développer une pratique de l'évaluation. L'évaluation de la qualité et de l'efficacité de l'aide devient, pour eux, un enjeu, bien que leur conception de cette qualité et de cette efficacité diverge de celle des donateurs occidentaux. Ils adaptent ainsi les « bonnes pratiques » d'évaluation des donateurs à leur propre culture, créant des pratiques « hybrides » d'évaluation. À travers l'exemple des donateurs émergents, nous voyons bien que l'évaluation s'inscrit « dans un ensemble de pratiques issues d'une trajectoire historique spécifique, de sorte que le foisonnement des règles, des lois et des pratiques permet que des choix soient opérés, que des aménagements soient faits et que des jeux entre normes et procédures contradictoires ou ambivalentes soient réalisés »[1124].

D'autre part, les donateurs du CAD-OCDE sont, à leur tour, influencés par les approches des donateurs émergents, comme nous l'avons vu dans le cadre du Haut Forum de Busan, et font évoluer leurs bonnes pratiques en les prenant en compte. Cette convergence pourrait mener,

[1124] HIBOU, Béatrice, 2012, *op. cit.* p. 166.

sur le moyen terme, à davantage de coopération tripartite entre les donateurs traditionnels, émergents et les pays bénéficiaires, pouvant mener à des exercices d'évaluations conjointes entre les deux groupes de donateurs[1125]. Ainsi, la fracture au niveau local sur la question de l'évaluation pourrait s'atténuer et l'on pourrait observer une convergence des pratiques parmi les donateurs. Cette perspective reste, toutefois, lointaine, dans la mesure où nous observons aujourd'hui que les donateurs traditionnels éprouvent déjà des difficultés à mener des évaluations conjointes entre eux.

[1125] Si les donateurs japonais et du nord de l'Europe sont impliqués dans le développement de coopération tripartie avec les donateurs émergents sur le terrain, les donateurs anglo-américains sont toutefois beaucoup plus réticents à ce sujet. Les donateurs français et allemand étant ouverts à cette idée, mais peu convaincu de la faisabilité de ce type de coopération en raison des contraintes de résultats et de standards environnementaux, sociaux, etc. qui pèsent sur eux. (Constat fait lors de participation à une réunion de travail de l'AFD et de ses partenaires du CAD-OCDE sur la coopération tripartite en décembre 2013).

Chapitre 8 : Les effets de l'évaluation au Sud, entre contrôle et démocratisation

Dans ce chapitre, nous allons analyser les effets qu'entraîne l'évaluation sur les relations entre donateurs et bénéficiaires. Nous analyserons si l'évaluation conduit, dans les faits, à une dépolitisation et à une « rationalisation » des relations entre donateurs et bénéficiaires. Nous étudierons également les effets de déclassement que produit l'évaluation en termes d'allocation de l'aide au développement. Nous vérifierons l'hypothèse selon laquelle les pays du Sud les plus « performants » seraient favorisés dans l'accès aux fonds d'APD. Nous nous demanderons ensuite si l'évaluation est un outil de domination sur les bénéficiaires de l'aide, comme le condamnent certains auteurs critiques, ou s'il est possible d'avoir une lecture plus complexe des rapports de force qui se nouent entre donateurs et bénéficiaires autour de l'enjeu de l'évaluation de l'efficacité des programmes d'aide.

A. Une allocation de l'aide davantage fondée sur la bonne gouvernance que sur la performance

Les travaux universitaires qui analysent les motivations de l'allocation de l'aide par les donateurs traitent peu des liens entre l'efficacité (comprise comme les résultats démontrés par les évaluations) de l'aide au développement et la répartition des fonds d'aide. La plupart de ces travaux se concentrent sur la mesure des liens entre les objectifs de politiques étrangères des pays occidentaux et la répartition géographique de leur allocation d'APD. Ils montrent, pour la plupart, que l'allocation de l'aide est motivée par des objectifs diplomatiques[1126]. En 2000, Jakob Svensson a examiné le lien entre l'allocation de l'aide et le respect des droits civils et politiques par les pays en développement, trouvant une corrélation variable entre ces deux critères selon les donateurs[1127]. En 2002, Alberto Alesina et Beatrice Weder se sont intéressés

[1126] On peut citer par exemple : GOUNDER, Rukami, "What Motivates Foreign Aid : A Case Study of Australia Aid to Indonesia", *Journal of Developing Areas,* vol. 33, n°3, 1999, pp. 379-394 ; ALESINA, Alberto, DOLLAR, David, "Who give Foreign Aid to Whom and Why ?", *Journal of Economic Growth,* vol. 5, 2000, pp. 33-63 ; NEUMAYER, Eric, *The Pattern of Aid Giving: The Impact of Good Governance on Development Assistance*, Routledge, 2003, 128p ; CANAVIRE, Gustavo *et al, Assessing the Allocation of Aid : Development Concerns and the Self-Interest Donors,* working paper n°1253, Kiel Institute for World Economics, 2005.

[1127] SVENSSON, Jakob, "Why Conditional Aid Does not Work and What Can Be Done About It?", *Journal of Development Economics,* vol. 70, n°2, 2000, pp. 381-402.

à l'effet de la corruption sur les montants d'aide au développement reçus[1128]. Leur étude ne trouve pas de lien significatif entre les deux phénomènes. Plus récemment, en 2008, une équipe de chercheurs de l'université d'Oxford a analysé, pour le compte du DIFD, les liens entre l'allocation de l'aide et trois critères : le besoin des populations, le mérite des gouvernements (niveau de démocratie et respect des Droits de l'Homme) et l'intérêt des donateurs[1129]. L'étude conclut que l'allocation de l'aide au développement est déterminée dans 58 % des cas par le besoin des pays, dans 38 % des cas par les intérêts politiques, économiques et historiques des donateurs et seulement dans 4 % des cas par la bonne gouvernance des gouvernements bénéficiaires. Elle tend donc à démontrer que les études précédentes auraient surestimé le poids des intérêts politiques par rapport à celui du besoin des pays dans les facteurs déterminant l'allocation de l'aide au développement. Le mérite, quant à lui, influencerait très peu cette décision.

Nous observons qu'aujourd'hui, un certain nombre de donateurs (la Banque mondiale et le DFID notamment) ont mis en place des indices de bonne gouvernance pour déterminer les montants et les destinations de leurs fonds d'aide au développement. Ces indices sont censés mesurer, de façon dépolitisée, le ratio entre la bonne gouvernance, le niveau de pauvreté des pays du Sud et les résultats des programmes d'aide, afin que les donateurs allouent leurs fonds de la manière la plus efficace possible. Ainsi, un pays très pauvre, avec une bonne gouvernance et dont les programmes d'aide sont évalués favorablement devrait être reclassé parmi les pays prioritaires de l'aide.

Ces indices ont avant tout vocation à donner une image plus juste et rationnelle de l'allocation de l'aide au développement. En effet, ils incarnent l'idée d'une aide allouée selon des critères scientifiques qui mesurent le volume d'aide que les pays « méritent » en fonction de leurs besoins et de leurs résultats.

Derrière la mise en place de tels indices « techniques », l'idée sous-jacente est que l'aide doit être allouée aux « bons » gouvernements du Sud, c'est-à-dire ceux qui méritent de recevoir de l'aide, car ils mettent en place des réformes permettant une amélioration de leur développement. Nous voyons ici que l'introduction d'indicateurs censés dépolitiser l'allocation de l'aide l'habille en réalité de tout un discours normatif. Cette dimension méritocratique de l'aide implique également que les pays les plus efficaces sont ceux qui reçoivent le plus d'aide

[1128] ALESINA, Alberto, WEDER, Beatrice, "Do Corrupt Governments Receive Less Foreign Aid?", *American Economic Review*, vol. 92, n°4, 2002, pp. 1126-1137.
[1129] HOEFFLER, Anke, OUTRAM, Verity, *Need, Merit or Self Interest – What Determines the Allocation of Aid,* Oxford University, Department of Economics, juillet 2008.

au développement. Critère de sélection qui apparaît souvent moins normatif que celui de la bonne gouvernance. Dès lors, l'évaluation *ex post* des résultats des programmes d'aide devrait devenir un élément direct de décision en matière d'allocation de l'aide, puisque c'est cet outil qui permet de déterminer l'efficacité selon laquelle les gouvernements mettent en œuvre les programmes d'aide, d'autant plus dans un contexte où l'aide budgétaire[1130] est une bonne pratique.

- **Des évaluations *ex post* peu prises en compte dans l'allocation de l'aide au développement**

Dans ce contexte, il nous apparaît intéressant d'étudier le poids qu'ont les résultats des pays bénéficiaires observés par les évaluations dans la décision de la répartition des fonds d'aide. L'efficacité des programmes de développement est-elle réellement devenue un facteur de sélectivité de l'aide ? Les évaluations *ex post* pèsent-elles dans la prise de décision sur la répartition des fonds d'aide ? À notre connaissance, un seul travail universitaire s'est intéressé à vérifier la corrélation entre l'évaluation de la performance de l'aide au développement et son allocation[1131]. L'étude en question porte sur 176 pays bénéficiaires et 20 donateurs bilatéraux. La méthode d'analyse présente, selon nous, quelques limites. L'indicateur utilisé pour mesurer la performance de l'aide est le *portfolio programme rating* de la Banque mondiale qui classe les pays bénéficiaires de la Banque à partir d'une évaluation réalisée en interne de l'efficacité des programmes d'aide par pays. Cet indicateur est transposé aux donateurs bilatéraux. Or, les programmes de ces donateurs ont des caractéristiques et des objectifs différents de ceux des banques de développement. Par ailleurs, il est important de souligner que la perception de l'efficacité des pays bénéficiaires peut varier d'un donateur à l'autre pour plusieurs raisons : l'ambition des objectifs donnés, l'expérience dans le pays, les relations politiques entre le donateur et le gouvernement partenaire, etc. Utiliser un indice prétendument universel d'efficacité des gouvernements bénéficiaires apparaît donc peu pertinent, dans la mesure où les donateurs n'ont pas de critères communs d'évaluation de cette efficacité. Par ailleurs, l'étude ne porte que sur des donateurs bilatéraux, alors que ce sont les organisations multilatérales qui développent le plus la pratique d'allocation de l'aide en fonction de la

[1130] Pour une définition de la notion d'aide budgétaire, voir lexique en annexe 7.
[1131] ISOPI, Alessia, MAVROTAS, George, *Aid Allocation and Aid Effectiveness, an Empirical Analysis*, United Nations University, UNU-WIDER, janvier 2006.

performance des pays bénéficiaires. Il apparaît donc primordial d'explorer les cas des banques de développement dans une étude portant sur ce thème.

Nous avons donc mené une analyse complémentaire pour tester la corrélation entre l'évaluation positive des résultats des programmes d'aide et les montants d'aide reçus par les pays bénéficiaires. Les cas d'étude que nous avons choisi de présenter sont : la Banque mondiale et le Fonds mondial de lutte contre le Sida, le paludisme et la tuberculose (FMSPT) pour les organisations multilatérales, et le Royaume-Uni et l'Australie pour les donateurs bilatéraux. Nous avons choisi ces donateurs, car ils sont un échantillon qui représente bien les acteurs qui produisent aujourd'hui des indices qui classent les pays bénéficiaires à partir d'évaluation de l'efficacité des activités d'aide. Nous utiliserons également les cas des États-Unis et de la France pour illustrer la manière dont la sélectivité est mise en œuvre par les donateurs qui n'ont pas recours à de tels indices. Notre analyse porte sur 176 pays bénéficiaires de l'aide. Au commencement de notre recherche, nous pensions que l'efficacité de l'aide, comprise comme les résultats des programmes d'aide au développement, serait un facteur déterminant de l'allocation de l'aide au développement, en raison de la hausse de la pratique de l'évaluation et de la rhétorique des bailleurs se vantant d'utiliser les preuves de l'évaluation dans leur prise de décision. Nous avons donc vérifié cette hypothèse.

Au début des années 2000, la Banque mondiale a créé le *Country Performance International Assessment* (CPIA) qui permet d'allouer l'aide en fonction de la bonne gouvernance des pays bénéficiaires décomposée en quatre critères[1132]. En 2010, la Banque mondiale a ajouté au CPIA un indice de mesure de la performance des projets d'aide : le *Project Portfolio Rating*. Ces deux indices forment le *Country Performance Rating* (CPR). Le CPR est largement repris par les banques de développement. Après avoir utilisé le CPIA pendant plusieurs années, le DFID a élaboré son propre indice d'allocation de l'aide, le *Need Effectiveness Index* (NEI). Il entend trouver le juste équilibre entre un fondement éthique de l'allocation de l'aide, basé sur les besoins des pays, et un fondement rationnel de l'aide, basé sur l'efficacité des interventions. Pour évaluer les besoins des pays du Sud, le NEI agrège les indices suivants : l'IDH du PNUD, l'indice de fragilité CIFP-FFS[1133] élaboré par l'université

[1132] Le CPIA classe les pays bénéficiaires de l'International Development Association selon 16 critères regroupés en quatre catégories : la gestion économique, les politiques structurelles, les politiques d'inclusion sociale et d'équité et la gestion du secteur public et des institutions.

[1133] Cet indice inclut des indicateurs d'autorité, de légitimité, de capacité, de gouvernance, de niveau économique, de sécurité, de développement humain, de démographie, d'environnement et de genre. Dans le calcul du NEI, le développement humain, le niveau économique et la démographie ont été exclus pour éviter une redondance avec les indicateurs de l'IDH.

de Carleton au Royaume-Uni et le seuil minimal de pauvreté (2 dollar par jour). Pour évaluer l'efficacité de l'aide, le NEI reprend l'indice CPR de la Banque mondiale.

Tout d'abord, nous allons analyser la corrélation entre le classement CPR de la Banque mondiale et les montants d'aide par habitant effectivement alloués aux pays éligibles à *l'International Development Association* (IDA). Prendre l'aide par habitant comme référence nous permet de comparer équitablement des pays ayant des densités de population différentes. Nos calculs ont porté sur les années 2005 à 2011, pour lesquelles les données nécessaires sont disponibles[1134]. Ils révèlent une corrélation positive entre les deux variables, dont le coefficient varie entre 0,4 et 0,5[1135].

Ces résultats montrent que la Banque mondiale alloue son aide de manière sélective en fonction de la bonne gouvernance et des résultats des pays du Sud. Parmi ces deux facteurs, nous nous sommes alors demandé lequel influençait le plus la politique d'allocation d'aide de la Banque mondiale. Pour répondre à cette interrogation, nous avons décomposé l'indice CPR selon les indicateurs qui le composent (voir ci-dessus). Nous avons ensuite mesuré la corrélation entre chacun de ces indicateurs et les allocations d'aide de la Banque mondiale.

Il apparaît alors que l'indicateur de résultat est faiblement corrélé aux montants d'aide par habitant, pour les années où les données sont disponibles, avec un coefficient de corrélation de l'ordre de 0,15. Nous en déduisons que les résultats d'évaluation des programmes d'aide n'ont que peu d'effets sur la prise de décision de la Banque mondiale en matière d'allocation de l'aide. En ce qui concerne les indicateurs CPIA de bonne gouvernance économique et de bonne gouvernance institutionnelle, le coefficient de corrélation est plus significatif ; respectivement de 0,3 à 0,4 et de 0,5 à 0,6.

Plus que l'évaluation *ex post* des résultats des programmes d'aide, c'est donc l'évaluation de la gouvernance des États qui sert d'instrument de sélectivité. Cette évaluation de la bonne gouvernance est de plus en plus étendue. Elle porte désormais sur l'ensemble de l'action des gouvernements, de la politique environnementale à la gestion des finances publiques, en passant par la garantie des égalités au sein de la population.

Ainsi, l'allocation est davantage déterminée en fonction de la bonne gouvernance que des résultats des programmes d'aide au développement. Cela reflète une tendance générale des donateurs à considérer que l'efficacité de l'aide est largement conditionnée à l'état de la

[1134] http://siteresources.worldbank.org/IDA/Resources/IDA16_Report-English-Final.pdf, consulté le 15 décembre 2014.
[1135] Les coefficients de corrélation de l'ensemble de notre analyse sont présentés dans un tableau récapitulatif en annexe 6.

gouvernance des États bénéficiaires. À ce sujet, Patrick Guillaumont écrit que « la performance est un concept ambigu. Objectivement, il réfère à la qualité des résultats obtenus par les politiques de développement ; mais dans les formules déterminant l'allocation de l'aide, il réfère à la prétendue qualité des politiques en elles-mêmes. Il semble que la signification de la performance est passée de l'efficacité de l'aide au développement à la bonne gouvernance des pays »[1136].

Les fonds verticaux[1137] utilisent également des indices pour affecter leur aide au développement. Le FMSTP adopte une approche sélective de l'aide fondée sur l'évaluation du besoin des pays (le taux de prévalence des maladies et le niveau de pauvreté des États) et de l'efficacité passée et prévue des programmes d'aide. Toutefois, ces indicateurs ne sont pas pondérés de manière équivalente. Le « besoin » et le « niveau de pauvreté » des pays comptent davantage que l'efficacité. Dans l'allocation des fonds d'aide aux pays, nous trouvons une corrélation très forte avec le taux de prévalence des trois pandémies dans les pays, ainsi qu'avec le niveau de pauvreté, alors que la corrélation est peu significative avec l'efficacité évaluée des programmes mis en place[1138]. Encore une fois, l'évaluation *ex post* de l'efficacité de l'aide ne semble pas être un facteur déterminant d'allocation.

Concernant les donateurs bilatéraux, le Royaume-Uni justifie et légitime son allocation de l'aide au développement par l'utilisation de l'indice NEI qui agrège des indicateurs d'efficacité et de besoin des pays du Sud[1139]. Il est frappant de voir que l'indicateur d'efficacité sur lequel se fonde l'indice NEI est l'indice CPIA de la Banque mondiale, c'est-à-dire une mesure de la bonne gouvernance des États. Ainsi, pour le DFID également, l'efficacité est assimilée à la bonne gouvernance. Le besoin des pays est calculé par leur niveau de pauvreté et leur niveau de fragilité et il est censé être davantage pondéré que l'efficacité.

Par la création de cet indice en 2011, l'objectif du DFID était d'afficher une meilleure redevabilité et une rationalisation de sa stratégie d'allocation de l'aide. Pour observer comment cet indice a modifié la politique d'allocation de l'aide du Royaume-Uni, nous avons suivi la même méthode de calcul que pour les organisations multilatérales. Nous avons mesuré la corrélation entre le classement NEI des pays et les montants d'aide au développement par

[1136] GUILLAUMONT, Patrick, *Adapting Aid Allocation Criteria to Development Goals,* An essay for the 2008 Development Cooperation Forum United Nations Economic and Social Council, mai 2008.
[1137] Pour une définition du terme fonds verticaux, voire lexique en annexe 7.
[1138] The Global Fund, *Overview of the Allocation Methodology (2014-2016), The Global Fund's new funding model,* mars 2014. Ces corrélations ont été établies dans le cadre d'une évaluation sur la coopération de la France en matière de santé qui est restée non publiée.
[1139] DFID, UKAID, *Bilateral Aid Review: Technical report,* March 2011.

habitant qu'ils reçoivent de la part du Royaume-Uni. Le coefficient de corrélation est globalement significatif, de l'ordre de 0,5. Nous notons une corrélation particulièrement forte concernant les pays en tête de classement NEI et recevant un haut volume d'aide par habitant.

Toutefois, il nous faut désormais vérifier si la création de l'indice NEI a effectivement conduit à rationaliser l'allocation de l'aide, entraînant un reclassement des pays allocataires de l'aide, ou bien s'il ne fait que refléter une politique qui préexistait à sa création. Nous avons donc comparé le classement des pays allocataires du DFID de 2011 et 2012 (2011 étant l'année de création de l'indice NEI) avec ceux des deux années précédentes. Nous obtenons alors un coefficient de corrélation très significatif, égal à plus de 0,9 entre les allocations d'aide de 2009, 2010, 2011 et 2012. Ce constat nous amène à conclure que la mise en place de l'indice NEI n'a aucunement modifié la politique d'allocation de l'aide du DFID par rapport à l'époque où il utilisait l'indice CPIA de la Banque mondiale. L'indice NEI a donc été utilisé comme un outil de légitimation d'une pratique de l'allocation qui ne correspondait pas à l'approche de la Banque mondiale. Le donateur ne fait pas correspondre sa politique aux indicateurs, mais les indicateurs à sa politique.

Reste à voir lequel des indicateurs de l'indice NEI pèse le plus sur l'allocation de l'aide au développement. Le coefficient de corrélation entre le volume d'aide reçu par habitant et l'indice CPIA n'est pas significatif pour les années 2011 et 2012 (ce qui est logique puisque ce facteur est moins pondéré que celui du besoin). Contrairement à la Banque mondiale, la bonne gouvernance, ou l'efficacité, puisqu'elles sont assimilées, n'est donc pas un facteur lourd de sélectivité. Parmi les indicateurs de besoin (pauvreté et fragilité), c'est celui de fragilité qui est le plus déterminant avec un coefficient de corrélation de 0,6. Nous en déduisons que l'allocation de l'aide du DFID est influencée par le niveau de fragilité des pays. Elle se dirige spécifiquement vers les pays en sortie de crise[1140].

L'Australie a également mis en place un indice pour évaluer la performance des pays bénéficiaires de son aide au développement et les classer en fonction, afin d'allouer de manière plus efficiente ses fonds d'aide au développement. L'efficacité est ici définie comme : « l'évaluation de la capacité des pays bénéficiaires à mettre en place de manière efficiente les programmes d'aide et à progresser vers des objectifs fixés en amont »[1141]. Les pays sont notés de 1 à 5 ; 1 catégorisant les pays « trop faibles » et 5 les « exemples de bonne pratique ». Le coefficient de corrélation entre les performances attribuées aux pays bénéficiaires et les

[1140] Les coefficients de corrélation de l'ensemble de notre analyse sont présentés dans un tableau récapitulatif en annexe 6.
[1141] AUSAID, *2013 Quality Review of Aid Program Performance Reports*, Office of Development Effectiveness, janvier 2014.

montants d'aide reçus est peu significatif (de l'ordre de 0,2). De la recherche d'autres facteurs de corrélation, il ressort clairement que ce n'est pas la performance des pays qui conditionne leur accès à l'aide australienne, mais leur appartenance à la zone d'influence géographique de l'Australie. En effet, lorsque nous comparons la situation géographique des pays et les montants d'aide, il apparaît clairement que l'aide australienne se concentre sur les pays de la zone Asie-Pacifique.

D'autres donateurs n'utilisent pas explicitement d'indices pour allouer leur aide au développement, comme USAID aux États-Unis ou les organismes français de développement.

Concernant la France, le ministère des affaires étrangères a publié une liste de 16 pays prioritaires[1142] sur lesquels l'aide au développement doit se concentrer. La cible privilégiée est les pays d'Afrique francophone qui figurent parmi les pays les moins avancés (PMA). La France lie donc son aide au développement à la promotion de la francophonie, à ses intérêts diplomatiques et au niveau de pauvreté des pays. L'évaluation de l'efficacité des programmes d'aide ne rentre pas dans ses critères d'allocation.

Nous avons ensuite cherché les facteurs qui pourraient peser sur les choix politiques d'allocation de l'aide d'USAID. Lorsque nous corrélons le classement des pays allocataires de l'aide des États-Unis avec un indice de gouvernance internationalement reconnu, le *World Gouvernance Index* de la Banque mondiale, nous obtenons un faible coefficient de corrélation, de l'ordre de 0,3 de l'année 2006 à 2011. Un facteur plus déterminant semble être la présence dans les pays d'une opération de maintien de la paix des Nations unies. En effet, lorsque nous corrélons les montants d'aide par habitant alloués par USAID et le budget des opérations de maintien de la paix dans les pays du Sud ou militaires, nous obtenons un fort coefficient de corrélation de l'ordre de 0,7. L'évaluation de l'efficacité des programmes d'aide ne rentre pas en compte dans l'allocation de l'aide d'USAID.

Cette analyse montre que les évaluations mises en œuvre pour juger de l'efficacité des programmes d'aide n'ont pas d'effets significatifs sur l'allocation de l'aide et ne sont pas un outil de dépolitisation des relations entre donateurs et bénéficiaires, tout du moins en ce qui concerne les montants d'aide alloués. Ce sont davantage les indices de bonne gouvernance ou de fragilité des États qui déterminent les stratégies d'allocation de l'aide des donateurs, soit des facteurs politiques et normatifs. Ainsi, ce n'est pas l'évaluation *ex post* des résultats des

[1142] Bénin, Burkina Faso, Burundi, République centrafricaine, Comores, République démocratique du Congo, Djibouti, Ghana, Guinée, Madagascar, Mali, Mauritanie, Niger, Tchad, Togo, Sénégal. http://www.diplomatie.gouv.fr/fr/politique-etrangere-de-la-france/aide-au-developpement-1060/dispositif-institutionnel-et/canaux-d-acheminement-de-l-aide/article/infographie-l-aide-au; consulté le 4 août 2014.

programmes d'aide, mais l'évaluation des politiques publiques des pays bénéficiaires qui a un poids important sur les pratiques d'allocation de l'aide des donateurs. Il s'agit en quelque sorte d'une évaluation *ex ante* de l'efficacité des programmes d'aide. Les donateurs prévoient que les activités d'aide seront efficaces en fonction de caractéristiques internes aux pays, évaluées en amont des dites activités. À partir de ces résultats, nous avons fait l'hypothèse que les évaluations de la gouvernance et/ou de la fragilité des pays bénéficiaires entraînent un effet de reclassement des pays allocataires de l'aide, entraînant l'éviction des plus faibles du régime de l'aide au développement.

- **Des populations stigmatisées du fait de leurs « mauvais » gouvernements**

Tout d'abord, nous constatons que les classements de « bons » et « mauvais » pays bénéficiaires varient significativement selon les bailleurs de fonds qui utilisent des indices d'évaluation différents. La variété des indices utilisés atténuerait donc les phénomènes de déclassement qu'ils pourraient entraîner pour certains États. En effet, les pays en queue de classement chez certains donateurs peuvent se retrouver en tête de classement chez d'autres. C'est par exemple le cas du Nigéria qui est le deuxième pays prioritaire de l'aide selon le classement du DFID, alors qu'il n'est que 45ème dans le classement de la Banque mondiale, ou encore de l'Éthiopie qui est troisième dans le classement du DFID et seulement 32ème dans celui de la Banque mondiale. Les pays bénéficiaires de l'aide tirent ainsi un avantage de ce manque de consensus entre les bailleurs de fonds pour maintenir le niveau d'aide qu'ils reçoivent.

Cette variabilité renforce encore notre constat selon lequel ces indices ne sont pas objectifs et neutres comme les donateurs le prétendent, mais sont bien le reflet de leurs orientations politiques. Nous avons vu que la création et la modification des indices à travers le temps ne modifient pas les pratiques d'allocation de l'aide. Par ailleurs, les indices de gouvernance mesurent des aspects sociaux et politiques (équité, fragilité des institutions, pouvoir de la société civile) en les traduisant en des indicateurs quantitatifs. Or, cette traduction et les opérations de pondération entre les indicateurs qui s'en suivent peuvent facilement être « instrumentalisées » par les donateurs, dans la mesure où la méthodologie utilisée est propre à chacun d'entre eux. De plus, la détermination des seuils de bonne gouvernance, de gouvernance acceptable et de mauvaise gouvernance varie pour chaque donateur. Dans sa thèse de Doctorat, Teresa Cravo démontre que la désignation par les donateurs de pays du Sud « bien » ou « mal » gouvernés, « efficaces » ou « inefficaces » est une construction cognitive reflétant davantage

des visions du monde subjectives que l'état réel de la gouvernance et de la performance des États. Elle prend comme cas d'étude le Mozambique, réputé être un « bon élève » en matière de gouvernance et de développement, ainsi que la Guinée-Bissau, stigmatisée par la communauté internationale comme étant un État retardataire dont la gouvernance est défaillante. Elle montre que le niveau de gouvernance de ces deux pays est en réalité similaire. Elle conclut alors que le *labelling* des pays du Sud est le fruit des relations de connivence qu'ils ont su, ou non, construire avec les donateurs[1143].

Ainsi, l'allocation de l'aide n'est pas rendue plus égalitaire par la création d'indices de performance et de gouvernance. Ces notions sont des construits subjectifs de donateurs influencés par leurs conceptions du « bon » parcours de développement et de la bonne gouvernance. Les États bénéficiaires sont toujours dans une position où ils doivent calquer leurs politiques publiques sur les attentes des donateurs. En effet, c'est plus la qualité procédurale des politiques publiques des pays bénéficiaires qui est évaluée que leurs résultats. La diffusion d'une culture de l'évaluation et du résultat chez les donateurs, n'a donc pas permis de dépolitiser l'allocation de l'aide, comme la littérature grise des organismes de développement tend à en diffuser l'idée[1144].

Par ailleurs, malgré la diversité des indices produits par les donateurs, lorsqu'une réputation de « mauvais » élève est trop attachée à un pays, il se trouve discriminé dans son accès à l'aide au développement. La marge de manœuvre qu'auraient les pays bénéficiaires pour maintenir leurs flux d'aide n'est donc tangible que dans les cas de pays entretenant des liens politiques et économiques étroits avec les donateurs, ces derniers pouvant alors passer outre des faiblesses de gouvernance. C'est ainsi qu'apparaît le phénomène des États « orphelins » de l'aide au développement, c'est-à-dire des pays qui sont délaissés par les donateurs du fait de leur « mauvaise » gouvernance. Au contraire, certains pays se retrouvent systématiquement reclassés comme pays prioritaires de l'aide au développement.

Le tableau ci-dessous présente les premiers pays bénéficiaires (aide par habitant) des donateurs membres du CAD-OCDE de 2005 à 2012 :

[1143] De ALMEIDA CRAVO, Teresa Paula, *What's in a label, the aid community's representation of success and failure in Mozambique and Guinea Bissau*, Department of Politics and International Studies, University of Cambridge, Supervisor Devon Curtis, janvier 2012, 271p.

[1144] European Commission, *Commission staff working document, EU Accountability Report 2011 on Financing for Development, Review of progress of the EU and its Member States*, 2011; DAC-OECD, *2011 OECD Report on Division of Labor: Addressing cross-country fragmentation of aid*, novembre 2011.

Tableau 3 : Les quinze premiers pays bénéficiaires de l'aide des pays du CAD-OCDE (2005-2012)

En MUSD	États	Moyenne aide reçue/hab./an
1	Cisjordanie et Gaza	552,75
2	Kosovo	368,5
3	Iraq	255,375
4	Timor-Leste	239,25
5	Afghanistan	188,875
6	Libéria	187,5
7	Bhoutan	165,125
8	Liban	157,5
9	Djibouti	157
10	Serbie	139,875
11	Congo, République du	139,125
12	Bosnie-Herzégovine	138,5
13	Géorgie	132,625
14	Nicaragua	126,375
15	Soudan du Sud	125,5
16	Haïti	124,625

Source : base de données de la Banque mondiale

Nous retrouvons majoritairement dans ce tableau deux types de pays : des pays en sortie de crise (Afghanistan, Irak, Palestine, Kosovo, Haïti, etc.) et/ou des pays à revenus intermédiaires ou en croissance rapide (Libéria, Géorgie, Nicaragua). L'aide se concentre sur un nombre restreint de pays, ces 16 pays recevant 25 % du total de l'aide par habitant fournie durant la période donnée. Ce tableau montre que le régime de l'aide actuel tend donc à ne pas prioriser la lutte contre la pauvreté, mais la stabilisation des institutions des Etats fragiles et la croissance économique. Ce phénomène peut s'expliquer par le fait que les pays en sortie de crise ou à revenus intermédiaires sont réputés comme ayant davantage de capacité d'absorption des fonds d'aide et comme offrant des résultats plus rapides. Ce phénomène est aussi révélateur d'un régime de l'aide au développement où les questions de sécurisation des États et de lutte contre la pauvreté sont de plus en plus liées[1145].

Pour pallier cette tendance, le CAD-OCDE attire depuis quelques années l'attention des donateurs sur les pays dits « orphelins de l'aide » qui se retrouvent déclassés dans le système de l'aide, notamment en raison de la hausse de la sélectivité basée sur une évaluation de

[1145] Pour une analyse du lien entre sécurité internationale et aide au développement, voir : CHATAIGNER, Jean-Marc, « aide publique au développement et réformes des systèmes de sécurité : l'improbable rencontre du Dr Jeckyll et de Mr Hyde », *Afrique contemporaine*, n°209, janvier 2004, pp. 39-49.

l'efficacité des gouvernements[1146]. Le tableau ci-dessous présente la liste de ces pays et leur degré de mise à l'écart de l'allocation de l'aide :

Tableau 4 : Pays orphelins de l'aide au développement (2013)

Pays	Aide insuffisante par rapport au nombre d'habitants	Aide insuffisante par rapport au PIB/hab.	Aide insuffisante par rapport à la bonne gouvernance	Allocation insuffisante par rapport à la performance	Nombre d'indicateurs
Madagascar	x	x	x	x	4
Malawi	x	x	x	x	4
Bengladesh	x		x	x	3
Gambie	x	x	x		3
Guinée	x	x	x		3
Niger	x	x		x	3
Togo	x	x	x		3
Népal	x		x		2
République Centrafricaine	x	x			2
Tchad	x	x			2
Comores	x	x			2
République démocratique du Congo	x	x			2
Érythrée	x	x			2
Guinée Bissau	x	x			2
Zimbabwe	x	x			2
Burkina Faso			x	x	2
Éthiopie			x	x	2
Sénégal			x	x	2
Tanzanie			x	x	2
Ouganda			x	x	2
Total 20 pays	**15**	**12**	**12**	**9**	

*Cet indicateur exprime le ratio entre le PIB/hab. et la population du pays
** Cet indicateur exprime le ratio entre le PIB/hab., la population du pays et sa bonne gouvernance
Source : CAD-OCDE, *Identification and Monitoring of Potentially Under-aided Countries*, 2013

Ce tableau identifie une vingtaine de pays « orphelins de l'aide », tous faisant partie des PMA. La sélectivité des donateurs conduit donc au déclassement d'un nombre significatif de PMA qui n'accèdent pas aux fonds dont ils auraient besoin pour faire vivre leurs populations dans des conditions décentes. Le tableau montre que ce déclassement est subjectif puisque

[1146] DAC-OECD, *Identification and Monitoring of Potentially Under-aided Countries*, 2013.

même des pays relativement bien gouvernés et efficaces sont exclus. Ce déclassement persiste sur de nombreuses années et il est difficile pour les pays de se reclasser[1147].

L'analyse développée dans cette partie montre que l'évaluation de l'efficacité des programmes d'aide n'est pas un facteur déterminant de l'allocation de l'aide au développement qui, si elle est prise en compte, l'est seulement à la marge. Au contraire, la gouvernance, qu'elle soit politique ou économique, est un facteur déterminant chez un certain nombre de donateurs, notamment les banques et les fonds multilatéraux de développement. Toutefois, chez la plupart des donateurs bilatéraux la corrélation entre évaluation de la gouvernance et allocation de l'aide reste peu significative. Pourtant, ces donateurs évaluent de plus en plus la gouvernance des pays en développement et construisent ou utilisent des indices pour classer les pays par types de gouvernance (institutionnelle, financière, sociétale, environnementale, etc.). Nous pouvons alors nous demander ce que motive cette volonté des donateurs d'évaluer la gouvernance des pays du Sud, si ce n'est la rationalisation de l'allocation de l'aide au développement.

B. L'évaluation, entre contrôle et démocratisation des gouvernements et des populations

En nous fondant sur les travaux d'auteurs ayant travaillé sur la sécurisation de l'aide au développement[1148] et d'auteurs critiques de l'évaluation[1149], nous suggérons que l'évaluation aurait un autre usage que celui de rationaliser des pratiques d'aide et de rendre plus égalitaires les relations entre les acteurs du développement. L'évaluation pourrait être utilisée comme un outil de contrôle des donateurs sur les gouvernements et les populations bénéficiaires dans un objectif de maintien de l'ordre international et de stabilisation des États fragiles[1150]. Cet usage de l'évaluation correspondrait bien à l'engouement des donateurs pour l'évaluation de la gouvernance des pays du Sud dans lesquels ils interviennent. Nous allons voir qu'à travers l'évaluation de programmes d'appui la gouvernance des États du Sud, les donateurs évaluent

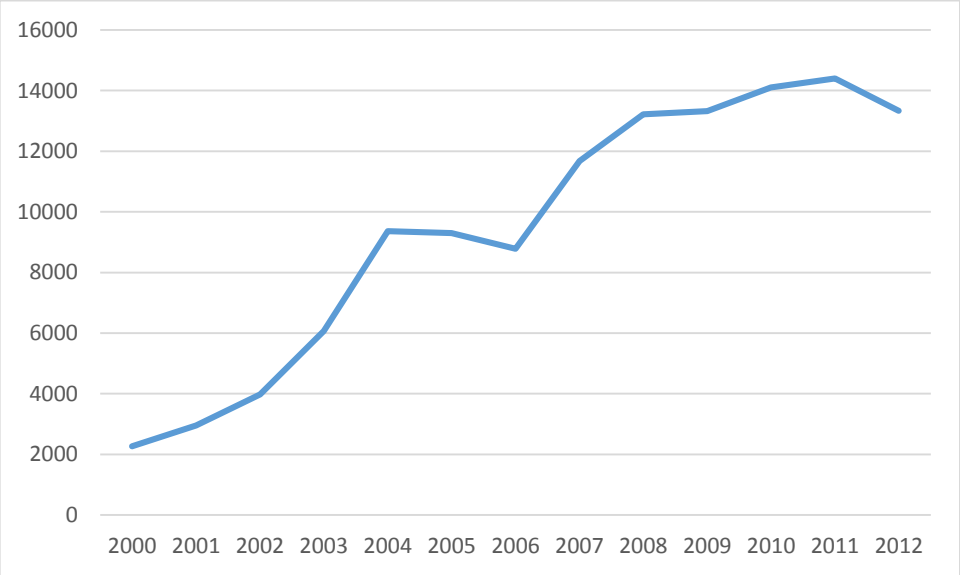
[1147] LAPORTE, 2011, *op cit.*
[1148] DUFFIELD, Mark, *Development, security and unending war: governing the world of people*, Cambridge: Policy, 2007, 266p.
[1149] GORI, Roland, del VOLGO, Marie-José, « l'idéologie de l'évaluation : un nouveau dispositif de servitude volontaire », *Nouvelle revue de psychologie*, vol. 2, n°8, 2009, pp. 11-26 ; ZARKA, Yves-Charles, l'Évaluation, un pouvoir supposé savoir, *Cités*, n°37, janvier 2009, 12p.
[1150] Pour une analyse du concept d'États fragiles voir notamment : CHATAIGNER, Jean-Marc, MAGRO Hervé, *États et sociétés fragiles : entre conflits, reconstruction et développement*, Éditions Karthala, 2007, 565p ; LAPORTE, Camille, 2011, *Évaluation et efficacité de l'aide au développement en Corée du Nord*, L'Harmattan, 2011, 220p.

en réalité davantage l'état de la gouvernance et les risques de déstabilisation de ces États que les résultats de leurs programmes.

- **Évaluer les programmes de renforcement de la gouvernance et de sécurisation des États**

Depuis le début des années 2000, la gouvernance est devenue une thématique importante de l'aide au développement, à laquelle des fonds croissants sont alloués par la majorité des donateurs. Des activités multiples sont regroupées sous la thématique de la gouvernance allant du renforcement des capacités de la société civile, à la lutte contre la corruption, ou encore à la mise en place de systèmes de suivi-évaluation dans l'administration. Ainsi, la gouvernance est un concept « fourre-tout » et évolutif, dont le contenu est difficile à cerner. Le graphique ci-dessous montre l'augmentation des actions d'aide qualifiée d'appui à la gouvernance de 2000 à 2012 :

Graphique 7 : Aide allouée à des projets d'amélioration de la gouvernance de 2000 à 2012 (MUSD)



Source : statistiques du CAD-OCDE

Selon nous, ce développement des actions visant l'amélioration de la gouvernance dans le cadre de la coopération internationale va de pair avec un phénomène de « sécurisation » de l'aide au développement. Nous entendons par là que les enjeux d'aide au développement et de sécurité internationale (maintien de la paix, lutte contre le terrorisme, etc.) sont de plus en plus liés par les donateurs. Les observations faites plus haut dans ce chapitre, démontrant une

corrélation, surtout vraie chez les donateurs anglo-américains, entre, d'une part la fragilité d'un pays ou la présence d'opérations de maintien de la paix des Nations-unies dans un pays et, d'autre part l'importance des volumes d'aide qu'il reçoit, vont dans le sens de cette hypothèse. Un autre signe de la volonté des donateurs de mieux intégrer le maintien de la paix, l'aide humanitaire et l'aide au développement est le débat actuel au sein du CAD-OCDE sur l'élargissement de la définition de l'APD pour y inclure des actions relevant aujourd'hui du maintien de la paix[1151]. Si les pays anglo-américains sont en faveur de cet élargissement, la France et l'Allemagne, en revanche, s'y opposent. Les premiers estiment que les enjeux de sous-développement (pauvreté, systèmes de santé lacunaires, faible niveau d'éducation, etc.) sont fortement liés aux situations de conflits sur lesquelles il faudrait agir pour permettre le développement, alors que les seconds craignent qu'en intégrant sécurité et aide au développement, l'objectif de réduction de la pauvreté soit encore plus mis à l'écart au profit de la stabilisation des institutions et du maintien de l'ordre

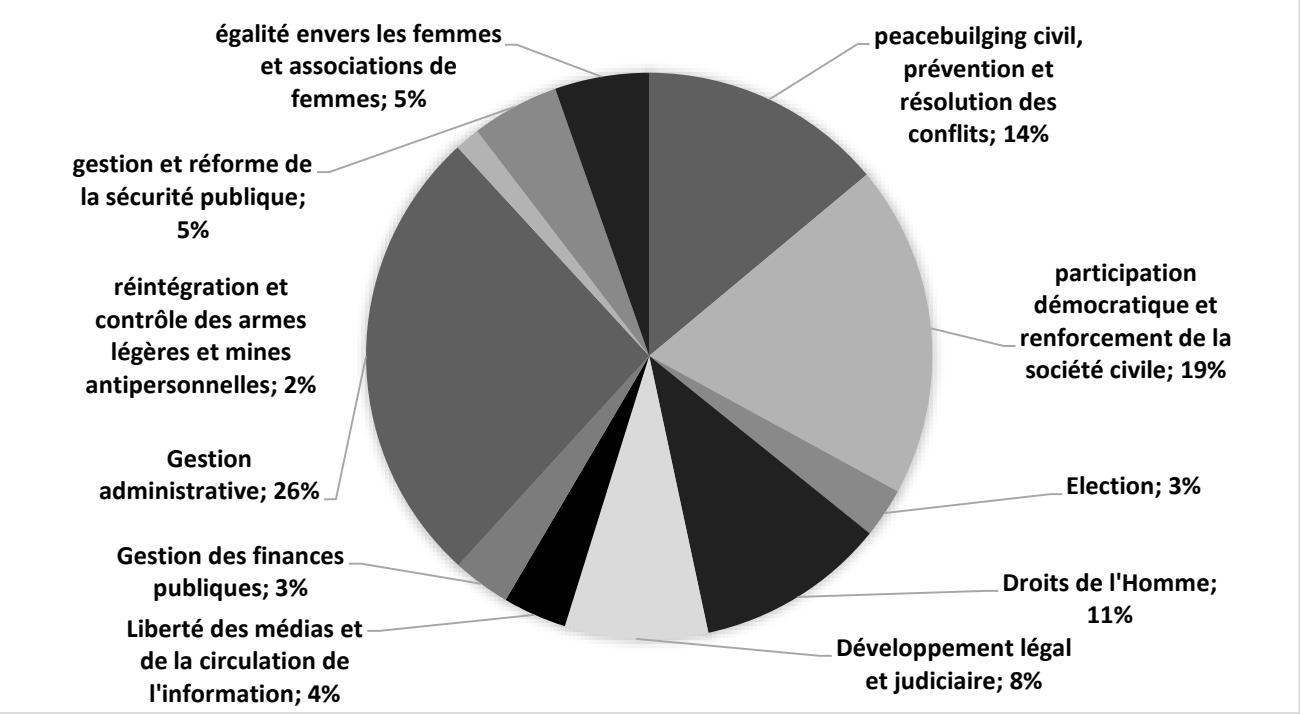
Les États-Unis et le Royaume-Uni lient de plus en plus leur aide au développement et leur politique de défense. Par exemple, le livre blanc du DFID, *Building our Common Future,* publié en 2009[1152], introduit l'idée d'une approche *peacebuilding-statebuilding* de l'aide au développement ; ou encore la *Strategic Defence Spending Review* de 2010[1153] souligne le rôle central de l'aide au développement dans la résolution des conflits et l'amélioration de la sécurité dans les États fragiles. De la même manière, l'USAID affiche clairement sur son site Internet qu'à « travers ses programmes d'aide au développement, l'USAID joue un rôle actif et fondamental dans la promotion des intérêts de la politique étrangère des États-Unis. Le développement a aujourd'hui sa place au côté de la défense et de la diplomatie, qui forment les trois composantes de la politique étrangère américaine »[1154]. Les organismes d'aide de ces deux pays font un plaidoyer auprès des autres donateurs pour diffuser l'idée d'une approche de l'aide au développement donnant une place centrale au renforcement des institutions et à la promotion de la paix et de la démocratie. L'ODI et le DFID ont ainsi publié plusieurs travaux de recherche démontrant les bénéfices d'une approche de l'aide au développement qui intègrerait des actions de sécurisation des États et d'amélioration de la gouvernance[1155].

[1151] OCDE, *Conflits et fragilité. Le Financement des États en transition : vers une meilleure réponse,* 2010.
[1152] DFID, *Eliminating World Poverty Building our Common Future,* White Paper, juillet 2009.
[1153] HM Government, *Securing Britain in an Age of Uncertainty, the Strategic Defense and Security Review,* octobre 2010.
[1154] http://www.usaid.gov/where-we-work ; consulté le 4 août 2014.
[1155] Par exemple, DFID, *Fighting poverty to* build a safer world, a strategy for security and development, mars 2005 ; WILD, Leni, ELHAWARY, Samir, *The UK's approach to linking development and security : assessing policy and practice,* ODI, Working paper 347, mai 2012.

Ces discours et travaux de recherche montrent que l'aide au développement est de plus en plus perçue comme un outil de stabilisation de l'ordre international. Selon Mark Duffield, plutôt que de réduire l'écart d'équité entre le Nord et le Sud, la coopération internationale a, aujourd'hui, pour objectif de contenir les effets déstabilisateurs du sous-développement[1156]. Selon lui, l'aide au développement serait conçue comme un « régime libéral de sécurité ».

À partir de la base de données du CAD-OCDE recensant les programmes d'aide au développement menés par les donateurs, nous avons calculé que, de 2008 à 2012[1157], environ 10 % des programmes d'aide au développement étaient consacrées à « l'amélioration de la gouvernance et au renforcement de la société civile » et à la « paix et à la sécurité ». Le graphique ci-dessous présente les activités incluses dans ces thématiques :

Graphique 8 : Programmes de soutien à la gouvernance



Source : base de données QWIDS du CAD-OCDE

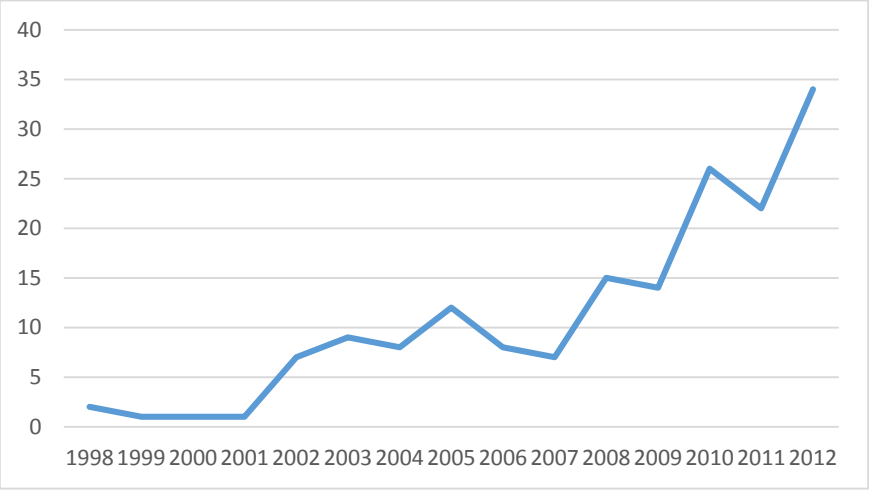
Cette représentation statistique montre qu'une part conséquente des activités liées à la bonne gouvernance a un lien direct avec la sécurité militaire (maintien de la paix, prévention des conflits, reconstruction des États, réintégration des enfants soldats ou contrôle des armes). Une autre catégorie d'actions porte sur l'amélioration de la gouvernance des États (renforcement des institutions légales et judiciaires, gestion des finances publiques, gestion

[1156] DUFFIELD, Mark, *Development, security and unending war: governing the world of people*, Cambridge: Policy, 2007, 266p, chapitre 1.
[1157] Dates pour lesquelles ces statistiques sont disponibles.

administrative, élections). Enfin, une dernière série d'actions se concentre sur la société civile (participation démocratique de la société civile, droits de l'Homme, liberté des médias et circulation de l'information, égalité hommes-femmes). Les programmes portant sur ces thèmes travaillent souvent à transformer les relations entre les États et leurs populations, et plus particulièrement à améliorer le statut social et la prise en charge des populations vulnérables (femmes, enfants soldats, minorités raciales et sexuelles, marginaux, etc.). Les donateurs qui ont le plus investi ce champ de l'aide au développement sont les États-Unis, le Royaume-Uni, l'Espagne, le Danemark, les Pays-Bas et les pays nordiques. Les donateurs comme la France, l'Allemagne, l'Italie ou les donateurs asiatiques se consacrent plus ponctuellement à ce type d'interventions[1158].

En conséquence de la multiplication des programmes d'aide portant sur la bonne gouvernance, les évaluations sur ce thème se sont multipliées depuis la fin des années 1990, comme le montre le graphique ci-dessous[1159] :

Graphique 9 : Évaluations des programmes d'aide touchant à la gouvernance



Source : DAC-OECD evaluation resource center http://www.oecd.org/derec/ consulté le 5 août 2014

Plusieurs questions se posent alors concernant la mise en œuvre de ces évaluations. Tout d'abord, est-il possible d'évaluer l'efficacité de programmes d'aide dont les résultats se révèlent sur le long terme et relèvent de changements immatériels et de long terme au sein des sociétés bénéficiaires ? Par exemple, il apparaît difficile d'évaluer à la fin d'un projet de deux ou trois

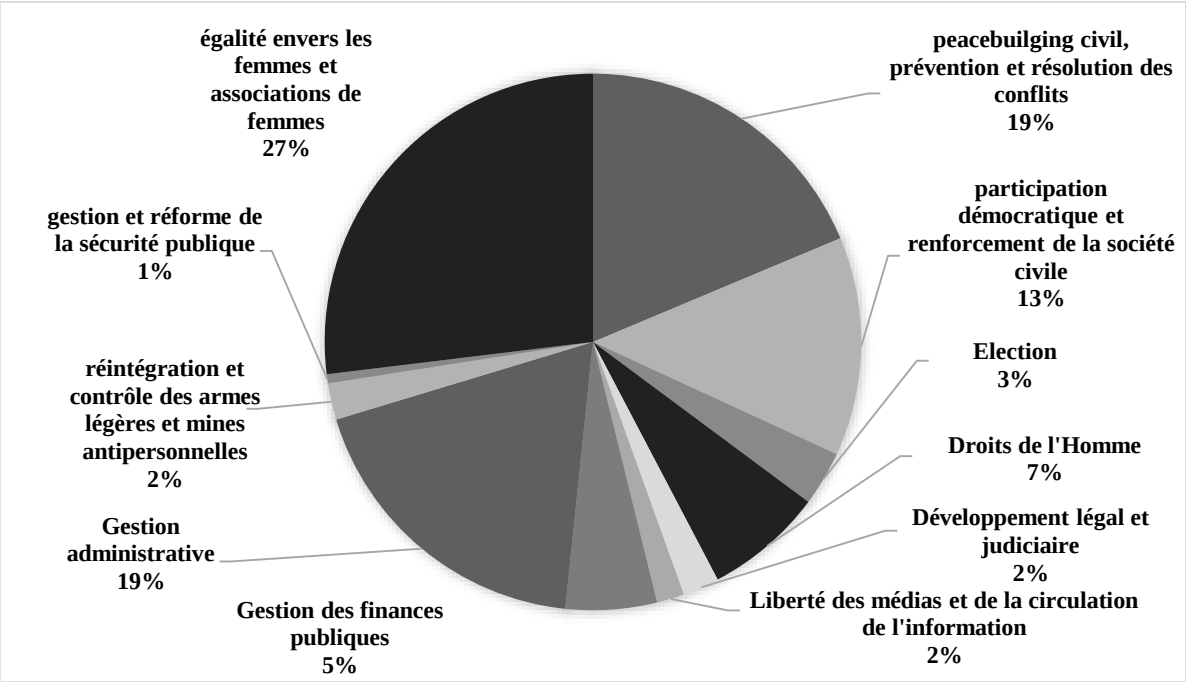
[1158] Entretien avec un membre du Centre for Development Policy Management, janvier 2010.
[1159] Nous ne prenons en compte que les évaluations de programme ou sectorielles car les évaluations de projets de sont pas recensées dans la base de donnée du CAD et ne sont pas rendues publiques par la plupart des donateurs.

ans la solidité d'institutions nouvelles, ou encore l'amélioration des conditions de vie de populations marginalisées et donc peu visibles.

Un autre questionnement concerne le rôle joué par l'évaluation dans la « sécurisation » des pays du Sud par les donateurs. En effet, si nous portons un regard foucaldien sur l'évaluation, celle-ci pourrait vue comme un outil de production de données quantitatives et des informations qualitatives sur des populations jugées « instables », ainsi que sur la qualité de la gouvernance des États pour « faire rentrer dans la norme » ou contrôler ces États et ces populations. L'évaluation permettrait ainsi aux donateurs de mieux contrôler l'environnement et la trajectoire des sociétés dans lesquelles ils interviennent. Nous observerions alors une tension entre le rôle affiché par les donateurs d'une évaluation servant un objectif de meilleure appropriation des programmes d'aide par les pays bénéficiaires et un objectif de contrôle de ces sociétés. Notre objectif sera de voir dans quelle mesure ces usages de l'évaluation s'observent sur le terrain, s'ils s'excluent l'un l'autre ou s'ils sont conciliables.

Afin de répondre à la première question, nous avons réalisé une revue des évaluations de programmes portant sur la gouvernance menées par les organismes d'aide de 2005 à 2014 :

Graphique 10 : Thématiques des programmes évalués



Source : DAC-OECD evaluation resource center http://www.oecd.org/derec/ consulté le 5 août 2014

Lorsque nous comparons le nombre de projets et programmes le nombre d'évaluations par thématique, nous constatons un décalage. Il apparaît que les programmes portant sur le

genre sont surreprésentés dans les évaluations, représentant 27 % d'entre elles, alors qu'ils ne constituent que 5 % de l'ensemble des activités de développement de ce secteur. Ce décalage pourrait s'expliquer par le fait que le genre est une thématique transversale qui est une composante de nombreux programmes de développement. Toutefois, cela est aussi vrai de la thématique des droits de l'Homme ou du renforcement des capacités de la société civile. Nous expliquons donc le grand nombre d'évaluations portant sur le genre par l'important plaidoyer des évaluateurs des organisations multilatérales pour en faire une thématique centrale de l'évaluation[1160]. Une autre raison relève du processus de normalisation dont a fait l'objet l'évaluation des problématiques de genre. Par exemple, les indicateurs élaborés par les unités d'évaluation des Nations unies pour mesurer l'égalité homme-femme dans des sociétés ont été repris par l'ensemble des organismes d'aide au développement et les guides de bonnes pratiques pour mieux traiter du genre dans les programmes d'aide se multiplient. Ces référentiels techniques sont moins développés dans les autres thématiques, telles que le soutien aux OSC, le renforcement des capacités judiciaires, le renforcement de la sécurité civile, etc., qui apparaissent quelque peu « sous-évaluées ». Nous voyons donc que plus une thématique du développement est normalisée, plus elle est évaluée.

Pour compléter cette analyse quantitative, nous avons porté le regard sur le contenu des évaluations, pour comprendre comment l'amélioration de la gouvernance y est appréhendée. Nous nous référons pour cette analyse à un échantillon d'une dizaine d'évaluations portant sur la démocratisation et l'amélioration de la gouvernance, effectuées ces cinq dernières années par des donateurs bilatéraux[1161]. Les objectifs de ces évaluations peuvent se résumer en quatre points :

- créer une compréhension commune des concepts de bonne gouvernance et de renforcement des capacités de la société civile ou des groupes vulnérables au sein des pays du Sud, en cartographiant et documentant les approches et les stratégies des donateurs dans des contextes variés ;
- tirer des leçons des interventions de renforcement de la bonne gouvernance, en identifiant ce qui a bien fonctionné et ce qui a moins bien fonctionné ;
- mesurer les effets des interventions de renforcement de la bonne gouvernance et analyser s'ils sont durables ;

[1160] *Cf.* Chapitre 4.
[1161] Ces évaluations sont consultables sur la base de données du CAD-OCDE rassemblant les évaluations menées par les donateurs membres : http://www.oecd.org/derec/publicationsdocuments/ consulté le 5 août 2014.

- explorer dans quelle mesure les interventions de renforcement de la gouvernance sont en accord avec l'agenda sur l'efficacité de l'aide.

Un premier point qui distingue ces évaluations de celles portant sur d'autres secteurs est qu'elles se consacrent en grande partie à l'analyse du contexte des sociétés dans lesquelles les programmes s'inscrivent, et notamment aux relations de pouvoir entre les groupes sociaux ciblés par les programmes. À la lecture de ces évaluations, il apparaît clairement qu'elles portent davantage un jugement sur l'état de la gouvernance des pays que sur l'efficacité des programmes d'aide. Les évaluateurs reconnaissent eux-mêmes ce biais : « il est de plus en plus reconnu que les résultats du soutien à la démocratie sont très dépendants du contexte d'intervention, que les données disponibles sur les programmes sont insuffisantes pour déterminer quels sont les facteurs qui mènent aux différents résultats et que les évaluations rigoureuses dans ce champ sont rares »[1162]. Ainsi, les évaluations ne parviennent généralement pas à mesurer les résultats des programmes d'aide et réalisent plutôt une critique des hypothèses, de la logique, des théories et des choix qui sous-tendent ces programmes. Les évaluations concluent souvent que les donateurs surestiment l'influence qu'ils peuvent avoir sur la gouvernance des pays bénéficiaires et qu'en conséquence, ils fixent des objectifs irréalistes, davantage fondés sur une conception théorique occidentale de ce que doit être la bonne gouvernance que sur le contexte et la culture politique des pays bénéficiaires. Cette critique des objectifs sert aux évaluateurs pour expliquer leur incapacité à appréhender l'efficacité de ces programmes. En effet, les évaluations ayant comme référentiel de jugement les objectifs fixés en amont des programmes, si ces derniers sont irréalistes, le jugement « objectif » de l'efficacité devient impossible.

Il existe, en conséquence, très peu d'évaluations qui analysent si les programmes portant sur l'amélioration de la gouvernance ont atteint les résultats escomptés. Celles qui s'essayent à cet exercice tendent à montrer que l'action des donateurs a des effets limités sur la démocratisation des États, l'amélioration de la gouvernance étant largement déterminée par des facteurs relevant du contexte historique et politique national. Les évaluateurs mettent en avant l'inadaptation de l'approche technocratique de la gouvernance des donateurs face à des enjeux fondamentalement politiques : « bien que le manque de compétences techniques et de capacités soit une contrainte significative à l'amélioration de la gouvernance, ce sont avant tout des facteurs politiques et culturels qui fondent le comportement des acteurs étatiques et non-étatiques. Ainsi, par exemple, les donateurs peuvent soutenir l'adoption de lois pour améliorer

[1162] NORAD, *Democracy Support through the United Nations,* Evaluation Department, octobre 2010, pp. 13-14.

l'égalité homme-femme, ou pour décentraliser le pouvoir gouvernemental, mais des blocages politiques pourront toujours empêcher la mise en œuvre de telles lois »[1163]. De plus, si certaines évaluations arrivent à démontrer des effets au niveau individuel, par exemple l'amélioration de la redevabilité de certains fonctionnaires ou le renforcement des capacités de groupes d'individus à s'organiser en OSC, ces résultats sont difficilement extrapolables à un niveau plus large.

Enfin, les évaluateurs ne parviennent pas à mettre en évidence la contribution des actions de renforcement de la bonne gouvernance à la baisse de la pauvreté ou à l'atteinte des OMD.

- **L'évaluation à des fins de gouvernementalité des populations des États du Sud**

Si les évaluations ne parviennent pas à atteindre leur objectif de mesure de l'efficacité des programmes de gouvernance et de sécurisation des États, elles apportent des éléments d'analyse intéressant sur l'état de gouvernance des pays du Sud. Ainsi, elles peuvent avoir une fonction implicite qui mérite d'être explorée : la « gouvernementalité ». Nous suggérons que, dans un contexte de « sécurisation » de l'aide au développement, l'évaluation serait perçue et utilisée par les donateurs comme un instrument de « gouvernementalité » des populations des pays bénéficiaires de l'aide, et ce au-delà des programmes centrés sur la bonne gouvernance. Nous nous référons à la définition donnée par Michel Foucault de la « gouvernementalité » : « l'ensemble constitué par les institutions, les procédures, analyses et réflexions, les calculs et les tactiques qui permettent d'exercer cette forme bien spécifique de pouvoir qui a pour cible la population, pour forme majeure de savoir l'économie politique, pour instrument technique essentiel les dispositifs de sécurité »[1164]. L'évaluation serait donc une procédure qui permettrait aux donateurs d'analyser, et d'exercer un pouvoir sur, les populations des pays bénéficiaires de l'aide, en guidant leur vie économico-politique et en sécurisant leur environnement et leurs comportements. D'aucuns pensent que l'évaluation serait une forme de « savoir » au service du « pouvoir » des donateurs[1165]. Nous voyons donc ici s'opposer deux conceptions de l'évaluation : celle d'une évaluation qui viserait l'autonomisation des individus et l'égalitarisation de l'aide au développement[1166] (évaluation participative, renforcement des capacités en évaluation, évaluation comme instrument d'appropriation de l'aide par les

[1163] *Ibid*. pp. 3-4.
[1164] FOUCAULT, Michel, *Sécurité, territoire, population: cours au Collège de France) 1977-1978*, Coll. «Hautes études». Paris: Gallimard: Seuil, 2004, 435 p, pp. 111-112.
[1165] FAURE, Alain, SMITH, Andy, « L'évaluation objet de savoir et enjeu de pouvoir », *Sciences de la société*, n°32, pp. 149-152 ; Zarka, Yves-Charles, 2009, *art. cit.*
[1166] *Cf.* Chapitre 7.

bénéficiaires), et celle d'une évaluation comme un « savoir-pouvoir » des donateurs qui renforcerait leur domination sur les gouvernements et les populations des pays du Sud (évaluation de la gouvernance, évaluation des risques, évaluation à des fins de régulation, etc.). Nous allons désormais analyser dans quelle mesure cette vision de l'évaluation comme un outil de domination est justifiée et s'il est possible d'avoir une lecture plus subtile du jeu d'acteurs entre donateurs et bénéficiaires qui se créent autour de l'enjeu de l'évaluation.

Comme nous l'avons vu préalablement, depuis le milieu des années 1990, la croyance prédomine parmi les donateurs que la gouvernance des pays influence largement l'efficacité des programmes d'aide[1167]. En témoigne, le recours croissant qu'ils font aux évaluations *ex ante* dans l'objectif de mesurer les risques et les opportunités des programmes d'aide en fonction des forces et des faiblesses de la gouvernance des États bénéficiaires[1168]. Pour ne pas prendre le risque d'initier un projet qui aura de mauvais résultats et donc qui nuira à la performance globale de l'organisme d'aide, il faut donc évaluer le plus finement possible, en amont, les caractéristiques de la gouvernance des États bénéficiaires. Or, l'observation participante que nous avons menée au sein de la division de l'évaluation de l'AFD, nous a permis de constater que l'efficacité de l'aide apparaît peu corrélée à la gouvernance des pays. En effet, les évaluations montrent que l'efficacité des activités d'aide varie fortement au sein d'un même pays, autant voire plus qu'entre différents pays. Cette variabilité de l'efficacité des programmes dans les pays se vérifie par une étude à plus large échelle des évaluations publiées par les donateurs du CAD-OCDE[1169]. Deux explications sont possibles : soit c'est la qualité variable des projets qui entraîne ces différences de résultats dans un même pays, voire dans un même secteur, soit le niveau contextuel qui influence l'efficacité des activités d'aide se situe au niveau des communautés ou des régions et non au niveau national,. La seconde explication est celle qui reçoit le plus de succès parmi les donateurs, peu enclins à remettre en cause la qualité de leurs programmes d'aide, ce qui les pousse à mener des évaluations *ex ante* de plus en plus fines et recentrées au niveau local. Cela justifie également le recours aux micro-évaluations expérimentales d'impact des projets.

[1167] *Cf.* Chapitre 3.

[1168] En témoigne l'importante littérature grise qui se développe sur le sujet des évaluations *ex ante*. Quelques exemples : OECD-DAC, *Promoting Pro-Poor Growth, Practical Guide to Ex Ante Poverty Impact Assessment,* 2007 ; BOURGUIGNON, François, FEREIRA, Francisco "Ex Ante Evaluation of Policy Reforms Using Behavioral Models", from *Tool Kit for Evaluating the Poverty and Distributional Impact of Economic Policies,* World Bank, 2003.

[1169] De récents travaux de recherche vont également dans le sens de ce constat, voir DENIZER, Cevdet, KAUFMANN, Daniel, KRAAY, Aart, « Good countries or good projects ? Macro and micro correlates of World Bank project performance », *Journal of Development Economics,* Vol. 105, novembre 2013, pp. 288-302.

Nous partons de l'hypothèse que l'évaluation de la gouvernance des États bénéficiaires de l'aide par un nombre croissant de donateurs et les classements qu'elle entraîne peuvent être assimilés à une forme de gouvernementalité. En effet, nous rapprochons cet exercice de celui de « normalisation » décrit par Michel Foucault comme « un repérage du normal et de l'anormal et l'élaboration de différentes courbes de normalité. L'opération de normalisation va consister à faire jouer les unes par rapport aux autres ces différentes distributions de normalité et à faire en sorte que les plus défavorables soient ramenées à celles qui sont plus favorables »[1170]. C'est bien là l'idée des donateurs qui réalisent ces évaluations de la gouvernance : établir une norme de bonne gouvernance et de stabilité politique dans les États du Sud et mettre en place des programmes d'aide au développement pour les aider à atteindre cette norme. C'est la première étape de la « gouvernementalité » : la mise en place d'un espace continu de différenciation, afin d'identifier les zones à hauts risques et celles à risques moins élevés. Selon Didier Bigo, il est possible de rapprocher cette démarche d'évaluation et de classification d'une « construction discursive de la menace » faite par des « agents sécurisateurs » qui participent à une nouvelle forme de « gouvernementalité », où savoir et pouvoir vont de pair pour légitimer de nouvelles lectures du monde portées par les États occidentaux. On se trouve dans un système international très hiérarchisé, qui donne plus de liberté aux acteurs du Nord pour intervenir dans les États fragiles[1171].

Selon Raphaëlle Parizet, « les savoirs experts du développement renforcent et figent des différences qui se construisent dans les espaces sociaux sans l'aide des chiffres et de la mise en forme savante : ils contribuent à maintenir des frontières cognitives et à réactiver des clivages. Outre les préjugés qui les accompagnent, ils entretiennent la domination des « développés » sur ceux stigmatisés par le « non-développement » ou le « retard »[1172]. En effet, en établissant des classements de la qualité de la gouvernance des États, les donateurs pointent du doigt « les mauvais élèves » et les stigmatisent aux yeux de la communauté internationale en remettant en cause leur légitimité à recevoir des fonds d'aide au développement. Les donateurs ont alors un pouvoir de contrôle et donc de domination sur ces États. En effet, les États ainsi stigmatisés, dépendant des ressources financières des donateurs et souhaitant améliorer leur image internationale, vont alors agréer aux exigences normatives des donateurs et entamer, au moins

[1170] FOUCAULT, 2004, *op. cit*, p.59.
[1171] BIGO, Didier, Sécurité et immigration: vers une gouvernementalité par l'inquiétude, *Cultures et Conflits*, Sécurité et Immigration, pp. 13-38, 1998, pp. 31-32.
[1172] PARIZET, Raphaëlle, « Mesurer le développement pour gouverner les peuples autochtones », *Revue Tiers Monde*, n°213, janvier 2013, pp.143-160

en apparence, des procédures de réformes de leurs politiques de développement, afin de se conformer aux critères de sélectivité des donateurs.

L'évaluation de la gouvernance des États ou l'évaluation de leur « mérite » à recevoir de l'aide au développement devient bien un instrument de pouvoir des donateurs sur les pays bénéficiaires[1173]. Pour contrecarrer cette critique, les donateurs avancent que ces évaluations de la gouvernance des pays du Sud se font en coopération avec les gouvernements de ces pays. Ils arguent que ces évaluations se fondent sur les « plans nationaux de développement », élaborés par les gouvernements des pays bénéficiaires de l'aide, pour déterminer le montant et les thématiques des programmes d'aide. Cependant, ces plans nationaux sont souvent réalisés par les gouvernements des États du Sud dans l'unique but de se conformer aux attentes des donateurs, sans répondre à une réelle réflexion stratégique. De l'aveu même d'un représentant du gouvernement d'un pays d'Afrique subsaharienne : « *des chiffres, des plans de développement, on vous en produit autant que vous en voulez, ils restent ensuite au fond d'un tiroir et personne ne les consulte jamais* »[1174]. Il est alors difficile de croire que ces documents traduisent réellement les besoins de la population de ces pays.

Les représentants des fonds verticaux, par exemple le FMSTP, se vantent particulièrement de sortir de ces relations de « savoir-pouvoir » autour de l'allocation de l'aide en demandant aux pays bénéficiaires de leur formuler eux même leurs demandes d'aide[1175]. Toutefois, il apparaît que ces procédures de formulation de demandes, où les pays doivent évaluer leurs besoins prioritaires et planifier des projets réalistes en fonction, sont difficiles à suivre par les pays les plus vulnérables qui disposent des capacités administratives les plus faibles[1176]. Ils se retrouvent alors discriminés dans l'accès aux financements. Nous voyons ainsi comment la technicisation des procédures d'allocation de l'aide au développement, sous des allures d'*empowerment* des pays bénéficiaires, consacre en réalité l'inégalité du régime de l'aide au développement. De plus, ce système introduit une forme de compétition entre les pays du Sud qui doivent séduire le donateur en lui proposant les projets qui apparaissent comme les plus rentables et efficaces.

[1173] GABAS, Jean-Jacques, « La mesure du développement : comment science et politique se conjuguent », *Revue Tiers Monde*, n° 213, janvier 2013, pp. 1-11.
[1174] Paroles recueillis lors d'un entretien dans le cadre d'une évaluation sur l'APD santé de la France auprès d'un officiel ivoirien travaillant dans le secteur de la santé.
[1175] Les pays sont regroupés dans différentes catégories selon leurs revenus, le niveau de prévalence des pandémies et leur « efficacité ». Chaque catégorie a le droit à une enveloppe globale dont la répartition entre les pays est fondée sur les demandes qu'ils formulent. Ces demandes sont évaluées par le fonds mondial qui juge de leur pertinence et de leur bien-fondé.
[1176] Entretien avec un haut fonctionnaire ivoirien dans le domaine de la santé publique, entretien à Abidjan, 26 mai 2014.

Après avoir ainsi « diagnostiqué » les États du Sud et leur avoir « administré » des programmes « correctifs d'aide, les donateurs suivent et évaluent l'évolution de leur situation par rapport aux standards de bonne gouvernance qu'ils ont fixés. Ce dispositif de suivi-évaluation produit un savoir sur les sociétés qui permet aux donateurs d'identifier les points de blocage et d'adapter en fonction leurs programmes. En effet, le suivi-évaluation va produire des données quantitatives et des informations qualitatives sur les populations des pays bénéficiaires de l'aide, à partir desquelles les donateurs vont pouvoir les « gouverner », c'est-à-dire se substituer à l'État défaillant pour réguler la société, et plus spécifiquement les comportements des populations les plus vulnérables ou « à risque »[1177].

Un autre signe qui amène à penser que l'évaluation devient un instrument de « gouvernementalité » est que sa forme expérimentale a une légitimité et une faveur grandissante auprès des donateurs. Or, les évaluations expérimentales se distinguent par la grande quantité de données statistiques qu'elles produisent sur les groupes de populations ciblés et par leur niveau « micro » d'application, puisqu'elles mesurent l'efficacité de l'aide au niveau des individus. Ces évaluations contiennent une phase de construction de données sur le comportement des individus. Les déviants sont ainsi identifiés et le projet d'aide tentera de les ramener dans la norme. Ainsi, l'aide au développement et l'évaluation s'individualisent de pair et suivent une logique foucaldienne de « normalisation ». Selon Rosalind Eyben, « *il est juste de parler d'utilisation de l'évaluation en terme de gouvernementalité sur les populations du Sud. Cette même utilisation est faite par les donateurs dans leurs propres pays sur les populations les plus vulnérables et marginales (les mères célibataires, les handicapés, les anciens prisonniers, etc.), c'est-à-dire celles qui sont vues comme des problèmes sociaux. L'évaluation expérimentale est utilisée pour gérer et réguler ces populations à risque* »[1178]. Il nous apparaît effectivement pertinent d'extrapoler cette analyse de l'utilisation des évaluations expérimentales au contexte des pays du Sud et de leurs populations vulnérables. En effet, une étude bibliographique sur les évaluations expérimentales menées par les donateurs, nous a permis de remarquer que les évaluations expérimentales visent le plus souvent des projets ciblant des populations vulnérables et ayant lieu dans les États fragiles[1179].

Grâce à l'évaluation et à tout un spectre d'instruments de « gouvernementalité » (production de bonnes pratiques, transferts de compétences, etc.), les donateurs vont exercer un

[1177] LAPORTE, Camille, 2011, *op. cit.*
[1178] Entretien, avec une ancienne employée au DFID, consultante pour des organisations de développement, 14 février 2014.
[1179] Étude réalisée à partir de la base de données DEREC du CAD-OCDE regroupant les rapports d'évaluation des donateurs membres, http://www.oecd.org/derec/publicationsdocuments/, consulté le 5 août 2014.

pouvoir « biopolitique » sur les populations des États dans lesquels ils interviennent. Toujours dans son analyse de la « gouvernementalité », Michel Foucault définit la biopolitique comme « un mode d'exercice spécifique du pouvoir qui cible les populations et cherche à en réguler les processus naturels tels que la mort, la naissance, les maladies, les circulations (humaines, de richesse, de marchandises, de savoirs etc.) et les risques que ces processus induisent, afin de générer une stabilité de long terme qui facilitera le développement des forces de l'État »[1180]. L'évaluation est un instrument de pouvoir, car elle porte un jugement qui se veut objectif sur les processus de développement, par rapport à une norme internationale que constituent aujourd'hui les OMD. Selon François Simonet, « si l'on évalue, c'est bien parce que cela permet de se situer par rapport aux systèmes de valeurs et de reproduire ou de produire des valeurs (…). L'évaluation répond à une conception du monde portée par l'évaluateur »[1181]. Évaluer suppose l'établissement d'une échelle de valeurs positives et négatives, à laquelle on confronte ensuite l'objet évalué. Les évaluations légitiment, sur une base d'autorité scientifique, des valeurs supposément objectives, qui en réalité, peuvent s'opposer à d'autres systèmes de valeurs. En plus de ce jugement, l'évaluation énonce des recommandations aux populations et aux structures qui les encadrent pour qu'elles se gouvernent de manière plus efficace. Mais ce jugement produit par les évaluateurs, ce pouvoir technique des donateurs, est-il utile aux gouvernements et aux populations ? Est-ce simplement un outil supplémentaire de domination du Nord sur le Sud, ou est-il possible d'avoir une lecture moins binaire, plus complexe, des effets de l'évaluation dans les rapports entre donateurs et bénéficiaires de l'aide ?

Deux lectures de cette utilisation de l'évaluation sont possibles. La première est négative et souligne une objectivation des bénéficiaires de l'aide qui sont considérés comme des « *problèmes à résoudre* »[1182]. Une évaluatrice du DFID développe, en ce sens, une analyse critique de l'utilisation de l'évaluation par cet organisme : « *l'approche du développement axée sur les résultats actuellement développée par le DFID est semblable à celle des bureaucraties coloniales où la catégorisation, la quantification et l'objectivation entraînèrent une manière de penser le développement qui considérait les individus comme des sujets objectivés nécessitant intervention et traitement. Ces individus sont ainsi réduits au silence et mis dans l'incapacité*

[1180] FOUCAULT, Michel et al, *Naissance de la biopolitique : cours au Collège de France (1978-1979).* Coll. «Hautes études». Paris : Gallimard: Seuil, 2004, 355 p, pp.185-187.
[1181] SIMONET, François, « L'évaluation : objet de standardisation des pratiques sociales », *Cités*, n°37, Janvier 2009, pp. 91-100, p. 95.
[1182] Entretien avec une ancienne Chief Social Development Advisor au DFID, consultante pour des organismes de développement, 14 février 2014.

de remettre en cause la vision qu'ont les donateurs des problèmes qui les affectent et de proposer des approches et des solutions alternatives »[1183].

Cette lecture postcoloniale, qui peut paraître extrême, est très présente aujourd'hui dans les travaux des anthropologues sur le développement[1184].

Les donateurs justifient cette objectivation des populations bénéficiaires et leur exclusion encore généralisée des exercices d'évaluation par plusieurs arguments, dont la trop faible capacité ou le manque de volonté des gouvernements et des populations à mener des exercices de suivi-évaluation, ou encore un souci de neutralité et d'objectivité des évaluations. Raphaëlle Parizet analyse ce mode de pensée des donateurs dans le cadre de programmes d'aide touchant aux peuples autochtones indiens : « à la mise à l'index des peuples autochtones dans les processus de fabrication des indicateurs, les agents des deux institutions (CDI et PNUD) apportent une justification en termes de compétence et de neutralité. Tout d'abord, selon eux, la technicité des processus de fabrication des indicateurs ne permet pas d'inclure les représentants autochtones dans ces processus, les réduisant à une supposée incapacité à y participer (…). Les agents ne considèrent pas l'exclusion des peuples autochtones de ces processus comme problématique dans la mesure où, selon eux, la neutralité et l'objectivité des savoirs experts garantissent et assurent leur impartialité »[1185]. Selon, un évaluateur de l'AFD, les évaluations expérimentales conduisent à nier le savoir local des populations du Sud en les objectivant et en leur imposant des recettes préconçues de « bon » développement[1186]. Nous retrouvons ici l'idée d'un retour à une conception scientifique de l'évaluation, porteuse d'une vérité sur des phénomènes sociaux et qui légitime le savoir qu'elle diffuse par la rigueur de sa méthodologie quantitative.

Nous proposons une deuxième lecture du jeu de « savoir-pouvoir » qui se construit autour de l'évaluation à partir de les notions de « gouvernement de soi » et de « souci de soi » théorisées par Michel Foucault. Le « souci de soi » est « l'obligation pour tout individu de se soucier de lui-même et qui est immédiatement liée, comme sa condition, à la connaissance de soi. On ne peut pas s'occuper de soi-même sans se connaître soi-même ». Foucault comprend le « souci de soi » comme la capacité de l'individu à s'occuper de lui-même, à se gouverner lui-même. Dans le chemin vers l'atteinte de la connaissance de soi, l'individu est à la fois sujet et acteur : « cet

[1183] Entretien avec une évaluatrice du DFID, 4 juillet 2014.
[1184] GEROMINI, Vincent, *Savoirs et politiques de développement : questions en débat à l'aube du XXIème siècle*, Karthala-Gemdev, 2008, 253p.
[1185] PARIZET, 2013, *art. cit.* p. 147.
[1186] Entretien avec un évaluateur de l'AFD spécialisé dans les méthodes expérimentales, 19 novembre 2013.

ajustement entre un gouvernement de soi qui se fera dans la forme de l'universel (comme discussion publique, raisonnement public et usage public de l'entendement) et, d'un autre côté ce qui va être l'obéissance, obéissance à laquelle seront contraints tous ceux qui font partie d'une société donnée, d'un État donné, d'une administration donnée »[1187]. Objectivation et autonomisation ne s'opposent donc pas inéluctablement. L'évaluation, lorsqu'elle se recentre au niveau de l'individu, si elle est réalisée de manière participative, peut constituer cet espace de discussion et de raisonnement publics. Elle peut apporter à l'individu cette connaissance de soi dont il a besoin pour se gouverner, et par extension gouverner les autres. Selon un ancien directeur du département de l'évaluation de l'AFD, « *c'est ici le véritable rôle de l'évaluation : échanger sur la façon dont les projets ont été menés, faire apparaître des consensus et des divergences. C'est l'étape de la maturité du projet, la dernière pièce de la coopération. C'est un moment privilégié qui permet de faire le bilan avec tous les acteurs et de construire un sentiment de confiance qui est primordial à la mise en œuvre efficace de l'aide* »[1188]. Si l'on considère cette approche de l'évaluation, elle devient alors, non pas un instrument de domination et d'objectivation des bénéficiaires de l'aide, mais un outil de leur autonomisation.

Si nous considérons le « pouvoir-savoir » exercé par les donateurs à travers l'évaluation comme « biopolitique », il peut être positif pour l'autonomisation des individus, dans la mesure où il aurait pour objectif de générer des forces au sein des sociétés, de les optimiser et de les ordonner. L'évaluation va produire des recommandations, répondant à des bonnes pratiques de gouvernement, auxquelles les individus seront incités à s'y soumettre afin de permettre l'ordre social, mais c'est là la condition du bon fonctionnement de toute société. De plus, les bénéficiaires de l'aide au développement gardent une relative autonomie dans l'application des recommandations produites par les évaluations. En effet, nous avons constaté que ces recommandations sont peu reprises dans les objectifs des nouveaux projets d'aide au développement et cette situation est parfois le fait de la résistance des administrations et des populations des pays bénéficiaires, lors des négociations de planification avec les donateurs, à les voir intégrées dans les projets[1189]. Il est donc réducteur de voir les objectifs des activités d'aide comme étant purement imposés aux pays bénéficiaires par les donateurs.

La mise en œuvre d'évaluations au niveau des populations vulnérables et marginalisées des pays du Sud peut aussi être vue comme un moyen de les rendre visibles, de donner une réalité à leurs problèmes qui n'en avaient pas auparavant, car ils n'étaient pas appréhendés. En

[1187] FOUCAULT, 2004, *op. cit*, p.37
[1188] Entretien avec un ancien responsable de l'unité d'évaluation de l'AFD, 21 décembre 2010.
[1189] Entretien avec une évaluatrice de la Banque mondiale, 20 décembre 2013.

effet, la bonne pratique voulant que toutes les évaluations soient publiées, la réalisation d'une évaluation entraîne une publicisation de l'enjeu évalué, d'autant plus lorsque la société civile, nationale ou internationale, s'empare des résultats des évaluations pour nourrir leur plaidoyer. Même si cette qualification par les évaluations des enjeux de développement auxquels se confrontent les populations vulnérables peut être jugée comme étant biaisée, car n'incluant pas assez leurs perspectives, elle a le mérite de les faire apparaître dans l'agenda politique international. Concernant la lutte contre le sida par exemple, des évaluations expérimentales ont montré que la stagnation observée dans la baisse de la pandémie était en partie due à la mauvaise prise en compte des populations vulnérables (homosexuels, travailleurs du sexe, toxicomanes) dans le dépistage des personnes affectées. La publicisation des résultats de ces évaluations a entraîné la mise en place d'actions spécifiquement ciblées sur ces populations et adaptées à leur mode de vie. Les évaluations ont servi de « preuves » pour appuyer le plaidoyer des ONG internationales de lutte contre le sida[1190].

Il apparaît donc limitatif de seulement penser l'évaluation comme un instrument visant à légitimer la domination des donateurs sur les bénéficiaires de l'aide. L'évaluation peut être un instrument de politisation des enjeux traités par l'aide au développement et un moyen d'émancipation et d'*empowerment* des populations.

Nous avons vu qu'en théorie, l'évaluation ne se limite pas à un instrument de domination des donateurs sur les bénéficiaires. L'évaluation peut contribuer à émanciper les populations vulnérables et à ouvrir un dialogue plus égalitaire entre donateur et gouvernement du Sud. Nous allons désormais étudier, de manière empirique, les effets du suivi-évaluation sur un gouvernement et des populations vulnérables du Sud à travers le cas d'un processus de suivi-évaluation des conséquences sociales d'un projet de grand barrage au Laos auquel nous avons pu participer[1191].

- **Évaluation et gouvernementalité : le cas d'étude des populations déplacées d'un barrage au Laos**

Le projet de grand barrage en question a provoqué le déplacement de 6 500 personnes autochtones, impact social important qui a été compensé par la mise en place d'un ensemble de

[1190] Exemple recueilli lors d'une évaluation menée pour le compte du ministère des affaires étrangères sur l'APD santé de la France durant des entretiens confidentiels avec les parties prenantes de cette politique.
[1191] Le processus de suivi-évaluation étant confidentiel, il nous est impossible d'apporter de trop nombreux détails sur ce projet. Le cas d'étude étant toutefois très intéressant nous avons tout de même choisi de l'exploiter dans la mesure du possible.

mesures économiques et sociales pour restaurer leur niveau de vie. Ce programme de compensation économique et sociale est intéressant dans le cadre de notre étude, car il a été accompagné d'une ingénierie très importante de suivi-évaluation. En effet, dans un souci de redevabilité, exacerbé par le caractère très politiquement sensible de ce type de projet de grand barrage, les donateurs ont souhaité se protéger de la critique des ONG internationales de protection de l'environnement en mettant en place un suivi étroit de la mise en œuvre de ces compensations et des évaluations *à* mi-parcours et *ex post* rigoureuses de leurs résultats[1192].

Deux objectifs ont été donnés au programme de compensation économique et sociale, à partir desquels il a été suivi et évalué :

- tous les foyers doivent voir leurs ressources augmenter et atteindre le seuil de pauvreté national cinq ans après le début de la période de réinstallation ;
- le niveau de vie des personnes réinstallées doit être « matériellement amélioré » (amélioration des conditions de santé, d'éducation, de logement, de mobilité) sur une base durable.

Comme nous l'avons vu précédemment, 150 indicateurs ont été mis en place pour évaluer l'atteinte de ces objectifs. Cinq panels de suivi-évaluation, internes et externes au projet, ont également été créés. Ces panels d'experts en sciences sociales (économie, sociologie, genre, anthropologie, environnement) ont été chargés de suivre et d'évaluer la mise en œuvre des compensations et de vérifier qu'elle allait bien dans le sens d'une augmentation du niveau de vie des populations déplacées.

Nous sommes donc ici dans le cadre d'un processus de suivi-évaluation qui a pour cible les individus et pour objet l'amélioration de leur niveau de vie. Or, cette notion « d'amélioration des niveaux de vie » est très subjective et relève de l'appréciation personnelle et communautaire des individus en question, qui peut être en décalage avec l'appréciation qu'en ont les donateurs. À la lecture des documents de cadrage du programme de compensations[1193], nous observons que l'amélioration du niveau de vie des populations a été restreinte à une dimension monétaire (augmentation du revenu des populations pour qu'il atteigne le seuil de pauvreté) et matérielle (accès aux infrastructures de bases dans les domaines de la santé, de l'éducation, du transport, de l'agriculture). Nous expliquons ce choix par la volonté des donateurs de pouvoir facilement évaluer quantitativement les progrès et ainsi les rendre publics pour contrer les critiques des ONG internationales de protection de l'environnement. L'évaluation est donc mise au service

[1192] Le projet en question a suscité une forte opposition de grandes ONG internationales de protection de l'environnement.
[1193] Document confidentiel.

des donateurs pour leur permettre de prouver, à court terme, qu'ils sont parvenus à améliorer visiblement le niveau de vie et la qualité de l'environnement des populations déplacées.

L'évaluation sert à valider et à légitimer les mesures de compensation mises en œuvre par les donateurs, en démontrant l'amélioration des conditions de vie monétaires et matérielles des populations et en omettant volontairement des critères, tels que la durabilité de la gestion des ressources naturelles, l'appropriation des populations bénéficiaires du nouveau mode de vie qui leur a été imposé, etc. Au lieu d'être un outil d'« émancipation » des populations vulnérables déplacées, le suivi-évaluation va participer à un processus plus global de « gouvernementalité » ou de « régulation » de ces populations. En effet, derrière les « améliorations techniques » que les donateurs proposent, se cachent des transformations politiques et culturelles lourdes imposées à la population, notamment la sédentarisation, l'arrêt de l'élevage des buffles, l'introduction d'une économie de marché, le changement des cultures agricoles et donc du régime alimentaire, etc. Le « savoir-pouvoir » des évaluateurs est mobilisé pour valider par un argumentaire technique ces transformations et pour les présenter comme « rationnelles » par rapport à un mode de vie jugé « ancestral ».

Cette instrumentalisation de l'évaluation à des fins de légitimation apparaît particulièrement lorsque nous observons la communication faite sur les résultats du programme de compensation. Dans les entretiens que nous avons menés, les donateurs et les opérateurs du programme soulignent que les populations déplacées sont « sorties de la pauvreté » grâce au programme de compensation et que leurs revenus ont augmenté significativement[1194]. Or, lorsque nous lisons les rapports des panels d'experts chargés du suivi-évaluation du programme, ces résultats apparaissent bien plus fragiles. Tout d'abord, les experts indiquent que les bases de données utilisées pour le suivi-évaluation comportent de nombreuses lacunes. Par exemple, elles ne prennent pas en compte l'augmentation de la population sur le site de réinstallation et ne s'intéressent pas aux inégalités entre les villages déplacés. Par ailleurs, la hausse des revenus semble significativement liée à l'exploitation illégale des ressources naturelles (pêche, forêt, etc.) de la zone de réinstallation et aux aides financières qu'ont reçues les populations à la suite du déplacement[1195]. La durabilité de cette prospérité économique apparaît donc fragile.

[1194] Propos tenus lors de réunions de travail dans le cadre de l'évaluation des impacts environnementaux et sociaux du barrage en juin 2012, voir liste des évènements observés en annexe 9

[1195] Constat tiré de nos discussions avec des donateurs et des membres de l'organisme opérateur de la construction du barrage en juin et juillet 2012 dans le cadre de l'évaluation des impacts environnementaux et sociaux du barrage, voir liste des évènements observés en annexe 9.

Nous remarquons également que le processus de suivi-évaluation a été mis en place de manière peu participative avec une inclusion limitée des autorités locales, elles-mêmes peu intéressées par cette tâche. Cette situation est le reflet de la difficulté générale qu'ont eu les donateurs et l'opérateur du projet à impliquer les partenaires laotiens dans le programme de réinstallation. En résultent des doutes importants sur la pérennisation des résultats du programme de compensation et sur l'effectivité du transfert de sa gestion et de ses structures communautaires et institutionnelles aux autorités locales une fois que les donateurs seront partis. Ainsi, il est simplificateur de parler d'une domination des donateurs sur les autorités locales, ces dernières s'étant peu préoccupées du sort des populations déplacées et s'étant peu intéressées à l'ingénierie sociale et environnementale déployée par les donateurs. Le gouvernement laotien était satisfait que les donateurs et l'opérateur du projet se substituent à lui pour « gouverner » les populations déplacées et les sédentariser.

Ce sont davantage les populations déplacées qui ont été objectivées par l'évaluation. Les populations vulnérables ont été considérées comme des « obstacles » à l'atteinte rapide des objectifs du programme qu'il fallait ramener dans la norme[1196]. Nous retrouvons donc bien une logique que l'on peut rapprocher de celle de la biopolitique, de la « gouvernementalité » et de la normalisation foucaldiennes. Si les populations ont été consultées sur le choix de la zone de réinstallation, les autres composantes économiques, agricoles, sociales, etc. du programme de compensation leur ont été plus ou moins imposées pour « le bien de leur développement », ou officieusement pour l'atteinte des objectifs du programme[1197]. Ce manque de participation dans le suivi évaluation est cohérent dans le cadre d'un programme qui a été mis en œuvre de manière peu participative. Le suivi-évaluation a alors perdu tout son potentiel d'instrument d'émancipation des populations qui apprendraient à « se gouverner elles-mêmes ».

Les témoignages d'évaluateurs que nous avons pu recueillir tendent à montrer que, de manière générale, les processus de suivi-évaluation intègrent peu les bénéficiaires finaux (les populations) de l'aide au développement[1198]. Les ajustements parfois opérés en cours de programme pour en améliorer l'efficacité sont le plus souvent décidés entre experts à la suite de recommandations d'équipes de suivi-évaluation ayant seulement consulté à la marge les bénéficiaires finaux. Nous le constatons dans notre cas d'étude où les ajustements, qui sont apparus nécessaires au vu, par exemple, de l'inadéquation flagrante du programme agricole

[1196] Documents confidentiels.
[1197] Constat tiré de nos discussions avec des populations déplacées par le barrage en juin et juillet 2012 dans le cadre de l'évaluation des impacts environnementaux et sociaux du barrage, voir liste des évènements observés en annexe 9.
[1198] Entretien avec une évaluatrice de l'AFD, 25 juin 2014 ; entretien avec un ancien évaluateur du DFID, 17 juin 2013.

initié au type de terre de la zone de réinstallation, ont été débattus entre les experts des organisations internationales de développement, sans prendre en compte les préférences et les capacités des populations pour tel ou tel type de cultures. En conséquence, les populations ont multiplié les stratégies de contournement des pratiques imposées par les donateurs et les résistances à la participation des enquêtes de suivi-évaluation puisqu'une partie de leur comportement ne se conformait pas aux procédures fixées par les donateurs.

Cette domination symbolique des donateurs sur les bénéficiaires n'est pas quelque chose de voulue. Sous une pression de résultats à court terme, les donateurs ont tendance à vouloir éluder les éventuelles résistances des populations du Sud aux recettes de développement qu'ils souhaitent voir appliquées. Il apparaît que plus les programmes d'aide sont politiquement sensibles, que ce soit vis-à-vis des partenaires locaux, ou vis-à-vis de « l'opinion publique internationale », comme cela a été le cas dans ce projet de barrage, plus le processus de suivi-évaluation s'inscrit dans une démarche de redevabilité et de production de preuves quantifiés de résultats. Les bénéficiaires de l'aide ont alors tendance à être davantage objectivés et les partenaires locaux tenus à l'écart, afin que les donateurs gardent le contrôle sur la mise en œuvre et les résultats du programme.

Par ailleurs, sans une volonté forte du gouvernement du pays bénéficiaire à s'impliquer dans le programme mis en œuvre (ce qui est rare dans les PMA), il semble peu probable que le suivi-évaluation puisse donner un caractère plus égalitaire à ses relations avec les donateurs, voire qu'il puisse atténuer cette distinction et la transformer en « partenariat », nouveau paradigme de l'aide au développement[1199]. Il existe des cas, surtout dans le cadre d'activités d'aide à destination de pays à revenus intermédiaires, où un gouvernement impliqué se positionnera comme un partenaire des donateurs dans la mise en œuvre des programmes d'aide et où le suivi-évaluation deviendra, en conséquence, un moment « catalyseur » de ce partenariat, où chacun expose ses attentes, ses approches et son expertise. La manière dont l'évaluation est réalisée reflète donc les relations politiques et les rapports de force préexistants entre donateurs et bénéficiaires, ainsi que de l'implication du gouvernement, dans les programmes mis en place.

Les effets d'exclusion des gouvernements du Sud que fait ressortir l'évaluation ne sont donc pas systématiquement une finalité délibérément poursuivie par les évaluateurs, mais peuvent être le résultat d'un désengagement des pays bénéficiaires des programmes d'aide, dû à leur désintérêt pour certains programmes ou à une volonté de se décharger de mesures

[1199] *Cf.* Chapitre 7.

impopulaires sur les donateurs. Dans tous les cas, pour que l'évaluation puisse accomplir son rôle d'égalitarisation des relations entre donateurs et bénéficiaires, ainsi que d'*empowerment* de ces derniers, il est nécessaire de renforcer le partenariat entre ces deux types d'acteur. À cet effet, le CAD-OCDE a développé la « bonne pratique » de redevabilité mutuelle cherchant à responsabiliser les donateurs et les bénéficiaires au sein d'un partenariat équitable pour l'atteinte d'un développement durable.

C. Remise en question de l'existence d'une « redevabilité mutuelle » entre donateurs et bénéficiaires de l'aide

Nous allons désormais étudier comment la notion de redevabilité mutuelle est appliquée et si elle permet aux pays du Sud de faire entendre leur voix auprès des donateurs, contribuant ainsi à une égalitarisation des relations dans le système de l'aide au développement.

▪ Une bonne pratique difficile à appréhender

La « redevabilité mutuelle » diffuse une image de l'évaluation comme un vecteur de « prise de pouvoir » des gouvernements bénéficiaires de l'aide par rapport aux donateurs. Elle a été développée par le CAD-OCDE au début des années 2000 et fait partie des cinq piliers de la Déclaration de Paris de 2005. Elle est définie comme « un processus par lequel deux ou plusieurs partenaires se tiennent responsables des engagements qu'ils ont volontairement pris l'un envers l'autre. Elle repose sur la confiance et le partenariat plutôt que sur des sanctions en cas de non-respect des engagements, afin d'encourager un comportement de changement nécessaire à l'atteinte des objectifs. La redevabilité mutuelle est fondée sur des données collectées et partagées entre les donateurs »[1200]. Ce concept vise à rompre avec une approche de l'aide fondée sur une redevabilité à sens unique des bénéficiaires de l'aide envers les donateurs, les premiers devant respecter les conditionnalités des seconds et fournir des preuves de leur progrès de développement, sans que les donateurs ne soient soumis à un quelconque contrôle par les gouvernements bénéficiaires sur la qualité de l'aide qu'ils fournissent.

Dans le cadre de la Déclaration de Paris, la « bonne pratique » de redevabilité mutuelle requiert des pays du Sud qu'ils i) renforcent, en tant que de besoin, le rôle des instances

[1200] DAC-OECD, *Mutual Accountability: an emerging good practice*, 2009.

parlementaires dans l'élaboration des stratégies nationales de développement et/ou des budgets ; ii) fassent davantage de place aux approches participatives en associant systématiquement un large éventail de partenaires à la formulation des stratégies nationales de développement et à l'évaluation de leur état d'avancement. Les donateurs, quant à eux, s'engagent à fournir en temps voulu des informations transparentes et détaillées sur l'aide fournie, afin de permettre aux autorités des pays partenaires de présenter des bilans et des perspectives budgétaires complets à leur parlement et à leurs citoyens. Les donateurs et les pays partenaires s'engagent à évaluer conjointement, au moyen des mécanismes objectifs existant à l'échelon local, les progrès qu'ils accomplissent respectivement dans l'exécution des engagements pris concernant l'efficacité de l'aide, notamment les engagements de partenariat[1201].

Nous observons donc que cette bonne pratique est étroitement liée à l'évaluation. Cette place centrale de l'évaluation transparaît également dans le fait que l'unique objectif et l'unique indicateur fixés à cette bonne pratique dans la Déclaration de Paris concernent l'évaluation. L'objectif à atteindre est que « tous les pays partenaires se soumettent à des évaluations mutuelles » et l'indicateur permettant de le mesurer : « le nombre de pays partenaires où sont entreprises des évaluations mutuelles des progrès accomplis dans l'exécution des engagements souscrits concernant l'efficacité de l'aide, notamment ceux qui figurent dans la présente Déclaration »[1202].

Ainsi, les donateurs voient en la mise en œuvre d'évaluations conjointes la solution pour responsabiliser les pays bénéficiaires et rendre plus égalitaires leurs relations avec ces derniers. La mise en œuvre d'évaluations mutuelles permettrait, aussi bien aux donateurs qu'aux pays bénéficiaires, de contrôler que leur partenaire respecte les engagements qu'il a pris au niveau international et au niveau local. Cette bonne pratique a été adoptée sous la pression de la « société civile internationale » et des pays du Sud qui demandaient un droit de regard sur les politiques d'aide des donateurs, jusqu'alors peu soumises à l'évaluation[1203]. Une représentante du CAD-OCDE déclare à ce sujet : « *la redevabilité mutuelle a permis de légitimer la demande des pays en développement d'avoir une aide qui répond à certains critères de qualité comme la prévisibilité ou l'alignement sur les priorités nationales* »[1204]. Cette affirmation sous-entend un changement majeur dans le système de l'aide : pour la première fois les bénéficiaires de

[1201] DAC-OECD, 2005, *op. cit.*
[1202] *Ibid.*
[1203] *Cf.* Chapitre 9.
[1204] Entretien avec une Senior Policy Advisor au sein du DAC-OCDE, 2 juillet 2014.

l'aide sont considérés comme des sujets, et non comme des objets de l'aide au développement, qui peuvent légitimement avoir des attentes quant à la manière dont les donateurs interviennent dans leurs pays et les faire valoir dans leurs dialogues avec ces derniers, aussi bien sur la scène internationale qu'au niveau local.

Si ce changement d'approche est un tournant important, la solution proposée par le CAD-OCDE pour rendre plus égalitaires les relations de l'aide au développement nous apparaît discutable. En effet, elle théorise qu'en augmentant les mécanismes de contrôle mutuel, en premier lieu desquels l'évaluation, entre les parties prenantes de l'aide au développement, leurs relations deviendront plus égalitaires. Le CAD-OCDE donne donc ici une solution technique à l'enjeu politique qu'est la transformation des rapports de force entre donateurs et bénéficiaires. Les acteurs sont considérés comme équitables lorsqu'ils possèdent des moyens comparables de contrôle les uns sur les autres.

Nous rapprochons la solution de la redevabilité mutuelle de l'idéologie du New Public Management qui imprègne la politique des donateurs depuis le milieu des années 1990 Dans son ouvrage *La bureaucratisation du monde à l'ère néolibérale*, Béatrice Hibou décrit une évolution globale des politiques publiques qui est semblable à ce que nous avons observé ci-dessus. Elle observe la concrétisation d'un contrôle de plus en plus bureaucratique dans les organisations publiques qui passe, non pas par la surveillance oculaire, mais par la mise en place de toute une série de normes et par la multiplication des procédures indirectes de contrôle du respect de ces normes. Selon elle, l'absence apparente de hiérarchie est concomitante de l'accroissement des processus de surveillance par les normes, ainsi que des procédures à respecter. Cela entraîne la suprématie de l'évaluation quantitative[1205].

Avec la redevabilité mutuelle, le contrôle « hiérarchique », qui s'exprime notamment par le fait que les donateurs imposent des réformes aux bénéficiaires, est remplacé par un contrôle sous la forme d'évaluations mutuelles. Le contrôle est donc partagé, dépersonnalisé et indirect ; la hiérarchie et la domination disparaissent en apparence. En contrepartie, les procédures à respecter (ici les « bonnes pratiques » de l'agenda sur l'efficacité de l'aide et du nouveau partenariat global pour le développement) se multiplient. Tout devient sujet à évaluation. Le contrôle et l'évaluation semblent être la contrepartie de la plus grande liberté donnée aux pays bénéficiaires pour gérer l'aide au développement qui leur est allouée, notamment avec l'aide budgétaire[1206]. Ainsi, nous identifions aujourd'hui cinq types de redevabilité :

[1205] HIBOU, 2011, *op. cit.*
[1206] Pour une définition de la notion d'aide budgétaire, voir lexique en annexe 7.

- la redevabilité des donateurs envers leurs tutelles financières et leurs contribuables, qui consiste à prouver la bonne utilisation des fonds publics d'aide au développement ;
- la redevabilité des gouvernements bénéficiaires envers les donateurs, qui consiste à démontrer le bon usage des fonds qu'ils ont touchés et les progrès de développement réalisés grâce à ces fonds ;
- la redevabilité entre pairs, selon laquelle des gouvernements bénéficiaires ou des groupes de donateurs s'engagent dans des processus de suivi-évaluation mutuels de leurs politiques d'aide au développement ou de développement ;
- la redevabilité des donateurs envers les gouvernements bénéficiaires, où les donateurs sont tenus responsables d'apporter une aide de qualité qui permette aux bénéficiaires d'améliorer leur état de développement ;
- la redevabilité des gouvernements bénéficiaires envers leur société civile et leur parlement, selon laquelle ils doivent être transparents sur les coûts, les modalités et les résultats des politiques publiques engagées grâce aux fonds d'aide internationale.

Les deux premiers types de redevabilité sont les plus anciens, ils se développent depuis les années 1960. Les trois derniers ont commencé à émerger dans les années 1990 pour devenir des enjeux forts pour les acteurs du développement au milieu des années 2000.

Nous retrouvons également la caractéristique de la suprématie de l'évaluation quantitative dans le régime de l'aide au développement. Par exemple, pour juger des progrès en termes de redevabilité mutuelle, c'est le nombre d'évaluations mutuelles réalisées qui est évaluée, alors que ce qui garantirait, selon nous, l'existence d'une véritable redevabilité mutuelle serait davantage la qualité ou le degré de partenariat des évaluations. Cependant, plus les objets à évaluer sont nombreux, plus il devient difficile de mener des évaluations qualitatives approfondies, auxquelles on préfère alors des évaluations quantitatives à partir d'indicateurs simples à renseigner.

Nous remarquons également que pour évaluer le respect de la redevabilité mutuelle, il est demandé aux acteurs de l'aide de réaliser une évaluation des évaluations qu'ils ont menées. Cette situation est une belle illustration d'un système où l'évaluation appelle l'évaluation. Selon Béatrice Hibou, le risque d'un tel système est une « dépersonnalisation des relations »[1207], où le dialogue entre acteurs se résume à un contrôle mutuel encadré par des normes, des bonnes

[1207] HIBOU, 2011, *op. cit.* p. 31.

pratiques et des indicateurs. Or, c'est précisément l'objectif des donateurs : dépersonnaliser les relations entre donateurs et bénéficiaires pour qu'elles apparaissent dépolitisées et dirigées par des procédures techniques et non par des considérations politiques.

Comme cela était prévisible au vu de la faiblesse des systèmes d'évaluation dans les pays du Sud[1208], la « bonne pratique » de redevabilité mutuelle, n'a guère été appliquée. L'évaluation de la Déclaration de Paris menée en 2011 qualifie la redevabilité mutuelle de « pilier orphelin » de l'agenda sur l'efficacité de l'aide[1209]. Les donateurs expriment des difficultés à conceptualiser et à mettre en œuvre ce concept. Par exemple, dans une note datant de 2008, l'agence australienne de développement Ausaid explique que « la redevabilité mutuelle est le principe de la Déclaration de Paris le moins compris, à la fois conceptuellement et opérationnellement. Le débat global sur l'efficacité de l'aide est toujours dominé par les donateurs. Les donateurs sont mal à l'aise avec l'idée de redevabilité mutuelle, selon laquelle ils devraient subir des sanctions lorsque leurs politiques de développement n'ont pas les résultats escomptés. En conséquence, cette bonne pratique a reçu peu d'attention des donateurs »[1210].

L'évaluation de la Déclaration de Paris estime que seulement 30 % des pays bénéficiaires de l'aide ont mis en place des systèmes de suivi-évaluation mutuels, dont la plupart ne sont pas effectifs[1211]. Nous partageons l'avis selon lequel cet objectif est peu pertinent. En effet, il y a peu de sens à accoler des mécanismes techniques de redevabilité mutuelle, tels que la réalisation conjointe de processus de suivi-évaluation, à des programmes d'aide dont l'élaboration et la gestion est assurée unilatéralement par les donateurs. La redevabilité mutuelle est un objectif politique et non technique, qui nécessiterait un engagement politique des donateurs et des partenaires à réformer l'aide, afin que la prise de décision soit égalitaire entre les deux parties prenantes. Sans une volonté politique de haut niveau et partagée, les mécanismes de redevabilité mutuelle resteront des coquilles vides.

Le Forum de Haut Niveau sur l'efficacité de l'aide de Busan était l'occasion pour le CAD-OCDE de donner un nouvel élan au principe de redevabilité mutuelle. L'objectif était de susciter un intérêt et un engagement politique des pays donateurs et bénéficiaires pour que cette bonne pratique soit mise en œuvre plus efficacement dans le futur. En conséquence, la redevabilité mutuelle est un axe important de l'accord de Busan. Cet accord stipule que les 150

[1208] *Cf.* Chapitre 5.
[1209] CAD-OCDE, 2011, *op. cit.*
[1210] Ausaid, *Mutual Accountability: « orphan » principle of the Paris Declaration,* ODE Issues Note, juin 2008.
[1211] DAC-OECD, 2011, *op. cit.*

signataires se tiennent « mutuellement redevables de la mise en œuvre de leurs actions respectives dans les pays en développement et à l'échelle internationale »[1212].

Au niveau des pays, les signataires se sont engagés à déployer des efforts pour mettre en place des cadres permettant le renforcement de la redevabilité mutuelle. L'accord prévoit également la mise en place d'un cadre international de suivi-évaluation mutuel des progrès des acteurs de l'aide vers une coopération plus efficace au service du développement[1213]. Bien que l'accord de Busan reprenne la « bonne pratique » de redevabilité mutuelle et la place en haut de l'agenda de réforme de l'aide au développement *post* 2011, nous constatons que les engagements pris sont moins contraignants que ceux de la Déclaration de Paris, où donateurs et pays partenaires s'étaient vus assigner des objectifs différenciés assortis d'indicateurs quantitatifs et qualitatifs. De plus, dans une note datant de 2012, le CAD-OCDE précise que « la participation aux efforts mondiaux pour effectuer le suivi de la mise en œuvre de l'accord de Busan est volontaire et n'est pas un préalable à la participation des parties prenantes au dialogue politique plus général »[1214]. Comme nous l'avons vu dans le chapitre précédent, cet engagement en demi-teinte est en partie dû à la volonté des membres du CAD-OCDE d'arriver à l'accord le plus inclusif possible, intégrant notamment les donateurs émergents peu enclins à s'engager au respect de procédures de redevabilité mutuelle avec les donateurs traditionnels. Selon une représentante du CAD-OCDE, « *à Busan, ce fut un succès d'amener la Chine et l'Inde, mais à quel coût ? La déclaration finale est inclusive, mais c'est un document vague aux objectifs peu clairs. Les pays en développement ont le sentiment que les donateurs traditionnels se désengagent, se défilent* »[1215]. Le caractère volontaire de l'engagement a inquiété les pays du Sud qui ont vu leur pouvoir de pression normative sur les donateurs, déjà faible, amoindri. En effet, sans engagement politique solide et contraignant de la part des donateurs, il apparaît peu probable que la bonne pratique soit davantage mise en place que durant les années précédentes.

Nous observons que, dans le cadre de la Déclaration de Paris, les donateurs ont eu tendance à assimiler redevabilité mutuelle, mesure des résultats et sélectivité de l'aide. Ainsi, les procédures de suivi-évaluation mises en place dans le cadre de la redevabilité mutuelle, obligeant les pays du Sud à fournir des données sur leurs progrès de développement aux donateurs, servaient à ces derniers à réaliser des évaluations *ex ante* des risques de leurs investissements dans les pays et des évaluations *ex post* de l'efficacité de leurs programmes

[1212] OECD-DAC, 2011, *op. cit.*
[1213] *Ibid.*
[1214] OECD/DCD/DAC, *Indicateurs, objectifs-cibles et processus proposés pour le suivi mondial dans le cadre du partenariat mondial pour une coopération efficace au service du développement,* Groupe de travail sur l'efficacité de l'aide, juin 2012.
[1215] Entretien avec une Senior Policy Advisor du DAC-OCDE, 2 juillet 2014.

d'aide, et non à l'instauration d'un dialogue. Nous estimons donc que la redevabilité mutuelle a été détournée de son principe originel (équilibrer davantage les rapports de pouvoir entre donateurs et bénéficiaires et améliorer le dialogue entre ces acteurs) pour servir des finalités de redevabilité, d'évaluation quantitative de l'efficacité de l'aide et de classement des pays du Sud selon leur niveau de gouvernance.

L'accord de Busan, portant le principe d'un « partenariat global pour le développement », a voulu rompre avec cette dérive et redonner à la redevabilité mutuelle sa vocation originelle. Il précise que les indicateurs créés dans le cadre de la redevabilité mutuelle « sont destinés à contribuer à un dialogue politique plus général sur la coopération pour le développement et à son efficacité, plutôt qu'à servir de tableaux de bord aux donateurs, au sens étroit, pour établir un classement des pays et des organisations »[1216]. Cette recommandation reconnaît implicitement que la redevabilité mutuelle a bel et bien été détournée par les donateurs dans un objectif de maîtrise des risques et de reclassement des pays bénéficiaires de l'aide. Nous notons ici une volonté du CAD-OCDE de distinguer clairement, d'une part la redevabilité unilatérale des donateurs envers leurs tutelles financières et leurs contribuables qui est axée sur l'évaluation des performances des programmes d'aide, et d'autre part la redevabilité mutuelle qui doit être un outil de dialogue et d'apprentissage.

Alors que jusqu'au Forum de Busan, les donateurs liaient la redevabilité mutuelle à la mise en œuvre d'évaluations quantitatives, nous observons aujourd'hui une évolution, avec un discours qui promeut davantage les bénéfices des méthodes qualitatives de suivi-évaluation. Dans une note datant de 2012, le CAD-OCDE déclare à propos de la redevabilité mutuelle que : « dans de nombreux cas, les méthodes qualitatives peuvent offrir le moyen le plus approprié (et parfois le seul) d'évaluer les progrès par rapport aux engagements de Busan. Les évaluations au niveau des pays, réalisées sous la forme de rapports narratifs produits par les pays en développement, seront dans le futur les méthodes d'analyse privilégiées. Les indicateurs devront s'inspirer de données factuelles de nature qualitative (…) qui transcendent l'accent plus étroit mis sur les indicateurs quantitatifs »[1217]. En effet, l'évaluation qualitative paraît être une méthodologie plus accessible et réaliste dans le cadre des pays du Sud où les données statistiques sont souvent très parcellaires. De plus, les méthodes qualitatives sont plus à même de donner la parole aux bénéficiaires de l'aide et d'ouvrir un dialogue entre les acteurs (par des ateliers de travail, des entretiens, des *focus groups,* etc.). Ces méthodes correspondent donc

[1216] OECD/DCD/DAC, 2012, *op. cit.*
[1217] *Ibid.*

mieux à la raison d'être de la redevabilité mutuelle, si on la considère comme étant plus large qu'un simple contrôle des acteurs les uns sur les autres.

Toutefois, l'exigence de mesure de la performance des programmes d'aide est toujours forte pour les organismes d'aide et ils ne pourraient pas utiliser ces méthodes qualitatives et les produits des évaluations mutuelles pour remplir leurs obligations de redevabilité nationale. Ainsi, la multiplication des niveaux de redevabilité implique une augmentation des processus d'évaluation. Les finalités des différents types d'évaluation et de redevabilité étant différentes, les rapports produits ne peuvent pas être mutualisés. Une évaluatrice du CAD-OCDE développe un argumentaire en ce sens : « *il ne faut pas être naïf, les bailleurs et les bénéficiaires n'ont pas les mêmes objectifs. Les donateurs doivent avant tout rendre des comptes à leurs tutelles financières et aux contribuables. Les pays bénéficiaires, quant à eux, veulent améliorer leurs politiques publiques. Ils ont un souci de redevabilité, mais il est plus faible et il n'est pas dirigé vers le même public. Sur certains points, les donateurs et leurs partenaires peuvent établir des analyses partagées, mais ils ne peuvent pas se rejoindre sur les conclusions, qui sont dépendantes de l'objectif donné à l'évaluation* »[1218]. Le développement de niveaux multiples de redevabilité entraîne donc une bureaucratisation de l'aide au développement, où les relations entre chaque catégorie d'acteurs sont régulées par des procédures diverses de suivi, d'évaluation quantitative, d'évaluation qualitative, de revues par les pairs, etc. chacun de ces termes exprimant une nuance de la notion générale de contrôle.

Par l'institutionnalisation de procédures de plus en plus nombreuses de redevabilité, les rapports de pouvoir du système de l'aide au développement ne sont pas rendus plus équitables, ils sont seulement moins visibles. L'objectif est de gouverner par la technique pour effacer le caractère politique des décisions prises, ce que traduit très bien le concept de « gouvernance »[1219].

Bien que le principe de redevabilité mutuelle ait du mal à se développer dans le système de l'aide au développement, certains pays bénéficiaires ont commencé à mettre en place des procédures pour évaluer la qualité de l'aide fournie par les donateurs. Ces exemples sont souvent utilisés par le CAD-OCDE comme des *success stories* pour démontrer l'évolution des pratiques de l'aide vers une approche plus partenariale avec les pays du Sud[1220]. Nous allons

[1218]Entretien avec la coordinatrice du réseau Evalnet au CAD-OCDE, 2 juillet 2014.
[1219] HIBOU, 2011, *op. cit.*
[1220] DAC-OECD, *Mutual Accountability: Emerging good practice*, 2009; DAC-OECD, *Ownership and Mutual Accountability, 10 top tips on mutual accountability*, 2010; ECA-OECD, *Development Effectiveness in Africa, Promise and performance: Applying mutual accountability*, octobre 2005.

désormais analyser en quoi ces dispositifs entraînent une redevabilité des donateurs envers les pays bénéficiaires. Notre objectif est de déterminer si ces dispositifs entraînent effectivement un *empowerment* des pays bénéficiaires et s'ils entraînent une hausse de la qualité de l'aide fournie par les donateurs.

- **L'évaluation des donateurs par les pays bénéficiaires : une illusion d'*empowerment***

À la suite de la Déclaration de Paris, des initiatives internationales, régionales et locales ont été mises en œuvre pour rendre les donateurs redevables envers les pays bénéficiaires de l'aide. Au niveau international, la « bonne pratique » de redevabilité mutuelle est incarnée par le processus de suivi-évaluation de la Déclaration de Paris qui a essayé d'intégrer au mieux les pays du Sud qui menaient leurs propres analyses de la qualité de l'aide dans leurs pays[1221].

Au niveau régional, le Nouveau Partenariat pour le Développement de l'Afrique (NEPAD), plateforme créée en 2001 pour rassembler les dirigeants africains, a pour but « d'établir de nouvelles relations de coopération (entre le Nord et l'Afrique) fondées sur le principe de redevabilité mutuelle vis-à-vis des principes de développement »[1222]. Dans ce cadre, le NEPAD a établi une collaboration avec le CAD-OCDE et institué un mécanisme de revue par les pairs permettant aux gouvernements membres de suivre et d'évaluer les « performances des donateurs par rapport aux engagements qu'ils ont pris sur la scène internationale »[1223]. Toutefois, selon le président de la société d'évaluation du Sénégal, « *si le NEPAD a mis en place tous les mécanismes nécessaires à cette tâche, sa faible influence politique et son manque de vision stratégique pour le développement de l'Afrique tendent à en faire une structure peu dynamique et peu apte à répondre à une ambition telle que la refonte des relations entre donateurs et bénéficiaires sur une base plus égalitaire* »[1224]. En conséquence, les rapports d'évaluation produits font plus état des points positifs et négatifs de la coopération pour le développement sur le continent africain qu'ils ne sont porteurs de recommandations novatrices pour améliorer les stratégies et les pratiques d'aide des donateurs suivant un principe d'équité et de partenariat[1225].

[1221] C*f.* chapitre 7.
[1222] Voir le site Internet du NEPAD : http://www.nepad.org/fr/about, consulté le 14 décembre 2014.
[1223] ECA/OCDE, 2005, *op. cit.*
[1224] Entretien avec le directeur d'un cabinet d'évaluation au Sénégal, président de la société sénégalaise d'évaluation, 3 avril 2014.
[1225] Haut Conseil de la Coopération Internationale, *Les priorités de la coopération pour l'Afrique subsaharienne et le nouveau partenariat pour le développement de l'Afrique (NEPAD),* Rapport, Avril 2002.

Il faut souligner que seul un petit nombre de pays du Sud ont mis en place des mécanismes pour évaluer les progrès des donateurs par rapport à leurs engagements internationaux et nationaux. Au dire même de certains hauts-fonctionnaires du siège du CAD-OCDE : « *sur le terrain le concept de redevabilité mutuelle est perçu comme un terme technocratique et pseudo-philosophique qui n'apportera que des procédures bureaucratiques supplémentaires* »[1226]. C'est un sentiment que nous avons effectivement retrouvé sur les terrains où nous avons pu nous rendre pour cette thèse (Laos, Côte d'Ivoire, Maroc). Toutefois, certains pays ont été innovants en la matière, se positionnant comme les « bons élèves » de la redevabilité mutuelle[1227].

La première catégorie d'États considérés (selon nous, de manière abusive) comme ayant mis en place des dispositifs de redevabilité mutuelle sont ceux ayant élaboré des adaptations locales de la Déclaration de Paris, comme le Laos ou le Viêtnam. Le CAD-OCDE vante les mérites de ces documents qui instituent des indicateurs et des objectifs cibles pour évaluer les pratiques et les résultats des donateurs et des gouvernements. Ils permettent aussi un dialogue continu entre les donateurs et les autorités locales, à travers des « groupes de consultation sur l'efficacité de l'aide », « des groupes de partenariat sur l'efficacité de l'aide », ou encore des « groupes de travail sectoriels ». L'évaluation des progrès vers l'atteinte des indicateurs cibles fixés par ces déclarations est censée être réalisée conjointement entre les donateurs et les pays partenaires. Le CAD-OCDE souligne également les progrès faits pour améliorer, dans ces pays, les bases de données statistiques sur l'aide au développement, pour renforcer le rôle de supervision de l'Assemblée nationale (Viêtnam) et pour améliorer la participation de la société civile dans la prise de décision politique.

Toutefois, notre enquête de terrain au Laos nous a amené à nuancer la capacité de ces versions locales de la Déclaration de Paris d'entraîner une redevabilité mutuelle entre les donateurs et les gouvernements des pays bénéficiaires. Au Laos, pays dont la gouvernance et les capacités des administrations sont faibles, les groupes de consultation entre les ministères et les donateurs sont largement dominés par l'expertise de ces derniers. L'élaboration de la version locale de la Déclaration de Paris a été un projet initié et dirigé par les donateurs. Enfin, le suivi des indicateurs établis revient en très grande partie aux donateurs puisque le gouvernement n'a que peu de capacités et d'enthousiasme en la matière. Les avancées en termes

[1226] Entretien avec une Senior Policy Advisor au CAD-OCDE, 2 juillet 2014.
[1227] DAC-OECD, *Mutual Accountability: Emerging good practice*, 2009; DAC-OECD, *Ownership and Mutual Accountability, 10 top tips on mutual accountability*, 2010; DAC-OECD, *Development Effectiveness in Africa, Promise and performance: Applying mutual accountability*, octobre 2005.

de redevabilité mutuelle se limitent donc à la création d'un espace de dialogue nouveau et ouvert à tous les acteurs de l'aide, même s'il n'est pas exploité à la hauteur de tout son potentiel, et à une meilleure transparence de l'aide, les donateurs s'engageant à rendre publiques les données sur l'aide qu'ils fournissent. Selon un représentant de l'AFD au Laos, « *il est difficile de parler aujourd'hui de redevabilité mutuelle, dans la mesure où le gouvernement du Laos n'a pas le poids et le pouvoir d'exercer une pression sur les donateurs pour influencer leurs politiques d'aide. Le gouvernement du Laos est fortement dépendant de l'aide internationale et a tendance à accepter toutes les formes d'aide qui lui sont proposées. Les procédures de redevabilité et d'évaluation qui retardent le déboursement des fonds d'aide sont même plutôt perçues négativement par le gouvernement, dont l'objectif principal est d'accéder le plus rapidement possible à des financements extérieurs* »[1228]. Ainsi, le gouvernement du Laos n'aurait ni la capacité, ni la volonté d'évaluer la qualité et les résultats des programmes d'aide gérés par les donateurs.

Nous estimons que pour qu'une redevabilité mutuelle émerge dans le cadre d'une relation d'aide au développement, il faut des gouvernements partenaires forts et relativement autonomes par rapport aux donateurs étrangers. Elle ne peut donc apparaître que dans le cadre de pays à revenus intermédiaires qui ont des politiques publiques solides et des marges de manœuvre vis-à-vis de l'aide au développement. À ces conditions, les pays du Sud peuvent avoir un pouvoir de négociation avec les donateurs et la capacité d'affirmer leurs positionnements politiques. Effectivement, nous constatons que la redevabilité mutuelle est bien plus effective au Viêtnam, pays à revenu intermédiaire, qu'au Laos. Par exemple, des OSC vietnamiennes, soutenues par le gouvernement, ont créé en 2009 leur propre plateforme de suivi-évaluation indépendant de la qualité de l'aide afin d'améliorer la redevabilité des donateurs et du gouvernement envers les citoyens vietnamiens[1229].

D'autres États cités dans les *success stories* des donateurs occidentaux sont ceux ayant mis en place des cadres gouvernementaux d'évaluation de la performance des donateurs. Le pays le plus cité en modèle, et le seul à avoir mis en place un tel cadre opérationnel, est le Mozambique[1230]. Le gouvernement du Mozambique évalue les performances des donateurs par rapport aux objectifs et indicateurs de la Déclaration de Paris et les classe selon les résultats atteints. L'ambition du gouvernement est de créer un phénomène de *name and shame* (nommer

[1228] Entretien avec un chargé de projets de l'AFD au Laos, 30 juin 2012.
[1229] GRAVES, Sue, *Mutual accountability at the country level, Vietnam country case study*, ODI-OECD, Joint Venture on Managing for Development Results, 2009.
[1230] SADEV, *Mutual accountability in practice, the case of Mozambique*, juin 2012.

et blâmer) pour inciter les donateurs à améliorer leurs pratiques d'aide. Cependant, les entretiens que nous avons menés tendent à montrer que les donateurs sont peu sensibles à ce classement effectué par le gouvernement du Mozambique[1231]. Les changements observés dans leurs pratiques à la suite de la mise en œuvre de ces exercices d'évaluation sont peu visibles.

En plus de ce mécanisme d'évaluation de la qualité de l'aide des donateurs, le gouvernement a mis en place un dispositif pour évaluer les résultats de son propre plan de réduction de la pauvreté. Il transmet les résultats aux donateurs afin de remplir sa part de redevabilité mutuelle. Parallèlement, les donateurs évaluent, de leur côté, les programmes d'aide qu'ils mettent en œuvre, ne pouvant, pour des raisons de redevabilité nationale, se reposer sur les évaluations réalisées par le Mozambique. Nous voyons donc ici concrètement comment le principe de redevabilité mutuelle implique une hausse des procédures de suivi-évaluation et donc une bureaucratisation de l'aide au développement.

Nos recherches ont montré que la redevabilité du gouvernement du Mozambique envers les donateurs prend largement le pas sur les autres formes de redevabilité. En effet, le gouvernement est redevable envers un grand nombre d'organismes d'aide au développement qui ne coordonnent pas leurs demandes de suivi-évaluation. En conséquence, les capacités humaines et financières en évaluation du gouvernement et des ministères sont absorbées par la redevabilité envers les donateurs, au détriment souvent de la redevabilité envers la société civile qui ne porte pas sur les mêmes attentes et qui est moins pressante. Les capacités du gouvernement d'évaluation de la performance des donateurs s'en trouvent également amoindries[1232].

Le Mozambique a aussi mis en place une base de données centralisant les informations transmises par tous les donateurs sur l'aide qu'ils fournissent. Pour améliorer le caractère participatif du suivi-évaluation du développement, le gouvernement a créé un « observatoire du développement » qui rassemble les donateurs, les représentants du gouvernement et la société civile en un forum participatif pour suivre la mise en œuvre du plan national de réduction de la pauvreté. Son objectif principal est d'intégrer la société civile dans la prise de décision politique en matière de développement économique et social. Des forums provinciaux ont aussi été créés[1233].

[1231] Entretien avec une évaluatrice du CAD-OCDE, 2 juillet 2014 ; entretien avec un analyste des politiques au sein du département de l'efficacité de l'aide du CAD-OCDE, 2 juillet 2014.
[1232] STEER, Liesbet, WATHNE, Cecilie, 2009, *op. cit.*
[1233] SADEV, 2012, *op. cit.*

Toutefois, en l'absence de statut légal de cet observatoire et d'un processus formel de suivi de la bonne application de ses recommandations, la capacité du laboratoire à influencer la politique du gouvernement est limitée. De plus, il faut noter que ce type de forums s'est multiplié au Mozambique ces dernières années surchargeant la société civile. La multitude des forums et des groupes de travail dans lesquels elle est impliquée limite son aptitude à tirer un avantage de ces opportunités. La mise en place de systèmes de redevabilité parallèles, internationaux et nationaux, peut ainsi nuire à la redevabilité du gouvernement envers la société civile, celle-ci se retrouvant happée par les exigences de la redevabilité mutuelle envers les donateurs.

Le Mozambique, le Rwanda et la Zambie sont les seuls pays à évaluer la performance des donateurs. Parmi ces pays, le dispositif du Mozambique est le plus élaboré. Les autres pays en développement se contentent de suivre et d'alimenter des bases de données sur l'aide fournie par les donateurs.

L'ensemble des cas que nous avons étudiés montre que la redevabilité mutuelle au niveau des pays est peu mise en œuvre. Les procédures de redevabilité des donateurs envers les gouvernements bénéficiaires sont des exceptions et ont encore peu d'influence sur la politique des donateurs. Nous trouvons plusieurs causes à ce blocage de l'application du principe de redevabilité mutuelle au niveau local :

- vu du terrain, cette bonne pratique est perçue comme étant très technocratique. Perception qui entraîne un sentiment négatif *a priori* des acteurs de l'aide au développement ;
- pour que le principe de la redevabilité mutuelle fonctionne, il faut qu'il y ait des incitations fortes et crédibles à respecter les engagements pris dans les accords nationaux entre donateurs et bénéficiaires. Dans la grande majorité des cas, les gouvernements partenaires n'ont pas la capacité de faire peser de sanctions sur les donateurs. Les incitations qu'ils font peser sont d'ordre relationnel ou réputationnel, ce qui est trop faible pour amener des changements de comportement. Ce déséquilibre entre les incitations des donateurs et celles des gouvernements bénéficiaires est inhérent aux relations d'aide au développement. Les donateurs contrôlent les fonds qu'ils allouent aux pays du Sud et peuvent les faire varier pour récompenser ou sanctionner les « bonnes » ou « mauvaises » performances des gouvernements du Sud. Ces derniers ne disposent pas de mesures équivalentes pour tenir les donateurs redevables de leurs performances. Toutefois, malgré cet avantage

comparatif que semblent posséder les donateurs, il est rare qu'ils retirent des fonds d'aide en raison de mauvaises performances des gouvernements récipiendaires. Il convient donc de relativiser le poids de l'évaluation des performances sur la répartition des fonds d'aide.

- l'usage de la redevabilité mutuelle à des fins de contrôle plutôt que d'apprentissage mutuel. La redevabilité mutuelle est utilisée par les donateurs comme un moyen de garder la main sur une aide qui suit de plus en plus la bonne pratique de l'allocation budgétaire[1234], tout en donnant l'image de laisser une plus grande liberté aux États bénéficiaires sur l'usage qu'ils font de l'aide qui leur est allouée.
- enfin, du dire même des représentants du CAD-OCDE, la redevabilité est perçue comme étant un idéal inaccessible, une illusion[1235]. Nous avons observé que les donateurs, quels qu'ils soient, estiment très étroit le recoupement entre les enjeux qui intéressent les donateurs et les gouvernements bénéficiaires. En résulte des mécanismes de redevabilité et d'évaluation parallèles, où ceux servant la redevabilité nationale des pays bénéficiaires et leur contrôle de la qualité des fonds d'aide sont défavorisés. Le fait que les donateurs voient peu de champs à évaluer conjointement avec les gouvernements bénéficiaires est révélateur d'un système de l'aide où l'évaluation est centrée sur les besoins des donateurs et non des pays bénéficiaires. Cela peut également laisser supposer que les programmes d'aide sont mis en œuvre parallèlement aux programmes de politiques publiques nationaux et qu'ils n'y sont pas intégrés. Même lorsque les programmes sont bien intégrés dans les programmes de développement nationaux, ce phénomène de systèmes parallèles d'évaluation se vérifie car les donateurs veulent évaluer leur propre contribution et les pays bénéficiaires refusent que les donateurs évaluent leurs politiques publiques nationales. Ainsi, de manière générale, ni les donateurs, ni les pays bénéficiaires ne souhaitent voir une intégration de l'évaluation de l'aide au développement et du développement ; pratique pourtant prônée le Partenariat de Busan pour le développement.

- **Conclusion de chapitre**

[1234] Contrairement à l'aide au projet traditionnelle, l'aide budgétaire n'a, quant à elle, pas d'objectif spécifique, identifiable. C'est le gouvernement du pays bénéficiaire qui décide, en compagnie des donateurs, quelle sera l'affectation des fonds.
[1235] Entretien avec une Senior Policy Advisor au sein du département de l'efficacité de l'aide au CAD-OCDE, 2 juillet 2014.

Alors que les donateurs développent une rhétorique importante sur *l'evidence based policy* et la gestion axées sur les résultats, ce n'est pas l'évaluation *ex post* de l'efficacité des programmes d'aide qui détermine aujourd'hui les allocations d'aide au développement, mais plutôt l'évaluation *ex ante* de la gouvernance des pays du Sud. Cette utilisation de l'évaluation entraîne un reclassement des pays prioritaires de l'aide. Ainsi, un certain nombre de PMA se retrouvent « orphelins de l'aide », c'est-à-dire évincés des financements des donateurs.

La pratique de l'évaluation devient donc de plus en plus normative. La maîtrise des risques étant un impératif grandissant que se donnent les donateurs, l'évaluation, et en premier lieu l'évaluation expérimentale, est aujourd'hui utilisée comme un instrument de « gouvernementalité » des populations du Sud les plus vulnérables. Le suivi-évaluation va contribuer à une entreprise de régulation des modes de vies de ces populations pour les faire rentrer dans une norme établie par des experts internationaux. Les gouvernements du Sud les plus fragiles se satisfont du fait que les donateurs se chargent de « gouverner » ces populations et se désengagent des programmes d'aide au développement et de l'ingénierie qu'ils mettent en place. Ainsi, les relations entre donateurs et gouvernements bénéficiaires sont plus de l'ordre de la substitution que de la domination.

Sur le terrain, le suivi-évaluation peut permettre un certain dialogue entre les donateurs, les bénéficiaires et les opérateurs de projets. Ainsi, le suivi-évaluation pourrait être un instrument d'apprentissage mutuel et un moyen d'objectiver les enjeux considérés comme importants par les populations du Sud sur la scène internationale. Toutefois, les donateurs étant soumis à une forte pression à donner la preuve de résultats à court termes, le but premier du suivi-évaluation devient de prouver que les populations ont bien atteints les objectifs qui leur avaient été fixés en amont des projets. La logique de contrôle et de normalisation prend alors le pas sur celle de l'apprentissage mutuel.

Conscients de cette limite de l'évaluation, les donateurs du CAD-OCDE ont mis en place la « bonne pratique » de redevabilité mutuelle, afin d'inciter les donateurs et les pays bénéficiaires à se tenir mutuellement responsables du développement dans le monde. Principe d'équité exemplaire en théorie, la redevabilité mutuelle reste peu appliquée et sa mise en œuvre diffère de la fonction pour laquelle elle a originellement été fondée. Nous observons ainsi comment l'instauration d'une bonne pratique d'après des considérations techniques peut être instrumentalisée lorsque le contexte politique dans lequel elle est mise en œuvre est dissonant par rapport à la finalité qui lui est donnée. Ainsi, la redevabilité mutuelle et les techniques d'évaluation qui la composent ont échoué à transformer les relations entre donateurs et

bénéficiaires. S'exprime donc ici une limite du « gouvernement par les instruments » de l'aide au développement. L'instrument de la redevabilité mutuelle n'est pas parvenu à renouveler les « rapports sociaux spécifiques entre la puissance publique et ses destinataires en fonction des représentations et des significations dont il est porteur »[1236]. Sans une volonté politique qui les porte, les « instruments » tels que l'évaluation n'ont ni de sens, ni d'utilité pour les acteurs et ont donc peu d'effets sur leurs pratiques.

[1236] LASCOUMES, Pierre, Le GALES, Patrick, 2009, *op. cit.*

Chapitre 9 : Les effets de reclassement de l'évaluation des donateurs

Le chapitre précédent a montré que les mécanismes de redevabilité mutuelle entre donateurs et bénéficiaires de l'aide n'ont que peu d'effets sur les politiques et les pratiques des donateurs. En effet, les pays du Sud manquent de moyens de pression pour inciter les donateurs à faire évoluer leurs approches de la coopération pour le développement. Cette forme de redevabilité, incarnée par la mise en œuvre d'évaluations partenariales avec les pays bénéficiaires et par la création de dispositifs permettant à ces derniers d'évaluer la qualité de l'aide des donateurs, reste aujourd'hui une vitrine servant à prouver à « l'opinion publique internationale » la bonne volonté des bailleurs de fonds à rendre plus égalitaires leurs relations avec leurs partenaires du Sud.

Ainsi, il semblerait que les donateurs ne sont soumis à aucune évaluation « externe » de leurs politiques d'aide au développement, au-delà des évaluations qu'ils commanditent eux-mêmes pour rendre compte de l'efficacité de leurs activités à leurs tutelles financières et aux contribuables de leur pays. Toutefois, depuis quelques années, nous voyons émerger de nombreux mécanismes d'évaluation « par les pairs » des politiques d'aide au développement des donateurs. À ces mécanismes s'ajoutent des évaluations produites par des centres de recherches indépendants. Nous nous sommes alors questionnée quant aux effets de ces évaluations sur les politiques et les pratiques des donateurs évalués. Quels donateurs se retrouvent déclassés par ces évaluations ? Les donateurs déclassés modifient-ils leurs manières de faire pour se conformer aux recommandations de leurs pairs ? Ces évaluations ont-elles un effet transformateur sur les politiques d'aide ?

Dans ce chapitre est analysée la façon dont des processus d'évaluation des donateurs par les donateurs se sont diffusés, à mesure que le régime de l'aide au développement s'est institutionnalisé avec l'élaboration d'accords normatifs engageant mutuellement les donateurs, tels que les OMD ou la Déclaration de Paris. Ces processus d'évaluation par les pairs ont des appellations multiples, notamment *reviews* (revues ou examens) ou *assessment*[1237]. Ils sont rarement qualifiés d'évaluation, bien qu'ils en possèdent toutes les caractéristiques, c'est-à-dire une analyse de l'efficacité, de l'efficience, de la pertinence, de l'utilité et, dans une moindre mesure, de l'impact de l'action des organisations d'aide au développement et de leurs résultats. Nous supposons que le terme « évaluation » est évité pour qualifier ces processus, car il

[1237] Terme n'ayant pas d'équivalent en français autre que le terme évaluation.

implique une notion de jugement de valeur sur les organisations évaluées qui est peu acceptable par les donateurs, encore guère enclins à être soumis à une redevabilité internationale.

L'évaluation par les pairs est un concept qui provient du domaine de la recherche académique. Il désigne le jugement critique et collectif d'un ensemble de chercheurs sur les travaux d'un ou d'autres chercheur(s), préalablement à leur publication ou à leur présentation dans des conférences. Selon Marcel Fournier, Yves Gingras et Creutzer Mathurin, « la notion d'évaluation par les pairs, chère au milieu académique, n'a de sens que lorsqu'il y a consensus sur la qualité des activités à évaluer, ce consensus n'étant que le résultat de la lutte pour l'imposition d'une définition de la recherche légitime »[1238]. En effet, si les « pairs » ne sont pas d'accord sur les critères déterminants la qualité du travail à évaluer, il est alors impossible qu'ils produisent un jugement collectif consensuel. La multiplication de ces processus d'évaluation par les pairs conduit donc à diffuser une conception unifiée de la recherche. Nous percevons ici comment derrière l'instrument technique de l'évaluation par les pairs se cache un processus politique de légitimation des approches dominantes et des acteurs qui les incarnent.

Le concept d'évaluation par les pairs a été repris dans le secteur de l'aide au développement dans un objectif d'apprentissage mutuel et de diffusion de la connaissance sur le développement[1239]. Nous démontrerons dans ce chapitre, que l'on retrouve dans les évaluations par les pairs appliquées au domaine de l'aide au développement, l'enjeu sous-jacent de légitimation des approches de l'aide portées par les donateurs qui réalisent ces évaluations. Nous démontrerons que les évaluations par les pairs sont le vecteur de diffusion d'une approche politique dominante de l'efficacité et de la qualité de l'aide. Nous montrerons que pour renforcer le pouvoir de changement de leurs travaux, les évaluateurs produisent des classements des donateurs, afin que ceux qui se trouvent déclassés soient incités à modifier leurs pratiques pour rentrer dans la norme. Ces évaluations créent ainsi une concurrence entre les donateurs, dans un système de l'aide où les acteurs se multiplient et où les fonds se restreignent.

A. Des politiques d'aide des donateurs bilatéraux encore opaques

Dans cette sous-partie, nous allons démontrer que, malgré une impression généralisée d'augmentation de la transparence dans le secteur de l'aide, les donateurs bilatéraux se

[1238] FOURNIER, Marcel, GINGRAS, Yves, MATHURIN, Creutzer, « l'Évaluation par les pairs et la recherche légitime », *Actes de la Recherche en sciences sociales*, n°19, septembre 1988, pp. 47-54, p. 54.
[1239] OCDE, 2008, *op. cit.*

soumettent encore peu à des exercices de redevabilité internationale[1240], gardant le contrôle sur l'évaluation et la communication de leurs résultats.

- **Des évaluations menées en cercle fermé**

Les « examens par les pairs » du CAD-OCDE ont été le premier mécanisme d'évaluation par les pairs entre donateurs. Ces examens ont débuté en 1962, soit la même année que la constitution du CAD-OCDE lui-même. Cet instrument est donc l'un des piliers du mandat du CAD-OCDE, qui compte parmi ses objectifs principaux l'échange de données quantitatives et d'informations qualitatives sur l'aide au développement, ainsi que l'apprentissage mutuel entre les donateurs. Les examens par les pairs ont été conçus pour identifier les bonnes pratiques mises en place par les donateurs et d'éventuelles difficultés communes, et ainsi pour ouvrir un dialogue et une coopération permettant l'amélioration de l'efficacité de l'aide au développement. Nous estimons, que par la création de ces examens par les pairs, un des objectifs sous-jacents du CAD a été d'entamer un processus d'harmonisation des pratiques de l'aide et de légitimation d'un ensemble de bonnes pratiques portées par les donateurs chefs de file du Comité, à savoir le Royaume-Unis et les États du Nord de l'Europe.

De par l'ancienneté des examens par les pairs, il est possible de faire une analyse de l'évolution de leur méthodologie, afin de déterminer si elle est en lien avec les changements des approches politiques de l'aide au développement. Chaque revue est préparée par une équipe composée de deux représentants du Secrétariat du CAD-OCDE et de deux représentants de pays membres. L'équipe mène un dialogue avec diverses parties prenantes du pays donateur évalué (agences de développement, ministères, parlementaires, acteurs de la société civile, etc.) et réalise des visites de terrain permettant d'appréhender plus directement la mise en œuvre des engagements du donateur dans les pays bénéficiaires[1241]. Les aspects évalués par ces examens ont largement évolué au fil des accords internationaux signés par les donateurs[1242].

Jusqu'à la fin des années 1990, les examens par les pairs étaient divisés en deux grandes parties. Tout d'abord, l'étude du cadre stratégique des politiques d'aide au développement qui évaluait la prise en compte des nouveaux enjeux de développement (environnement, bonne

[1240] Par redevabilité internationale, nous entendons des initiatives d'évaluation des politiques d'aide des donateurs menés par des structures d'évaluation étrangères aux pays desdits donateurs. Ces structures peuvent être d'autres donateurs dans le cas d'évaluation par les pairs, des centres de recherche, ou encore des ONG.
[1241] Entretien auprès de personnels de l'OCDE chargés d'encadrer les revues par les pairs, 4 juillet 2014, Paris, Siège de l'OCDE.
[1242] Notre analyse se base sur l'étude d'une trentaine de revues par les pairs réalisées entre 1980 et 2014.

gouvernance, genre, etc.) par les donateurs. Une courte analyse de l'efficacité de la gestion de l'aide était également réalisée, qui se concentrait principalement sur les processus d'allocation de l'aide au développement (volume d'aide, répartition géographique, répartition sectorielle, part d'aide liée, etc.). Les relations des organismes d'aide avec les autres acteurs du développement, à savoir les ONG, les organisations multilatérales, le secteur privé et les pays en développement, étaient également brièvement abordées.

Les évaluations se concentraient donc sur l'efficacité organisationnelle des organismes d'aide et sur l'analyse de leurs processus d'allocation des fonds et de mise en œuvre de leurs activités. L'efficacité de l'aide sur le terrain et les résultats des activités n'étaient pas abordés, ce qui en faisait des exercices peu utiles aux acteurs extérieurs aux agences d'aide bilatérales. Nous allons désormais montrer que, dans les années 2000, les examens par les pairs se sont davantage apparentés, dans leur contenu, à des exercices d'évaluation.

Au début des années 2000, la structure des rapports est restée la même, mais l'évaluation du cadre stratégique des donateurs s'est désormais concentrée sur les efforts mis en œuvre pour contribuer à l'atteinte des OMD à l'horizon 2015. Nous constatons également que le contenu des examens par les pairs s'est étoffé au fil des ans et couvre un nombre d'aspects de plus en plus large.

À partir de 2005, les examens par les pairs se sont concentrés sur l'évaluation de la mise en œuvre des « bonnes pratiques » de la Déclaration de Paris. Les thématiques de la gestion de l'aide dans les États fragiles, de la prévisibilité de l'aide, du soutien politique à l'aide au développement, etc. ont également été ajoutées aux examens. Cette superposition d'anciens et de nouveaux principes à évaluer laisse penser que les donateurs ont du mal à se réformer et à atteindre les bonnes pratiques qu'ils s'engagent à respecter, même plusieurs décennies après leur adoption. Elle montre aussi que de plus en plus d'aspects de l'aide bilatérale sont soumis à une exigence de transparence et de redevabilité internationale[1243], même si cette dernière est longtemps restée limité au cercle des « pairs » des donateurs.

À la suite du forum de Busan et de l'adoption du Partenariat global pour le développement, les examens par les pairs ont changé de perspective et ont mis davantage l'accent sur les enjeux de l'élaboration d'une politique globale du développement intégrant les autres dimensions de politique étrangère (sécurité, économie, diplomatie) des donateurs et prenant mieux en compte les attentes des pays bénéficiaires.

[1243] Pour une définition de cette notion, voir chapitre 8.

Les recommandations produites par les examens par les pairs vont toutes dans le même sens, quelle que soit l'organisation étudiée[1244]. Elles contribuent donc bien à l'émergence d'un consensus parmi les donateurs occidentaux sur ce qu'est une aide au développement de qualité. Pour les donateurs bilatéraux, recevoir une bonne évaluation par les examens par les pairs signifie que leurs pratiques sont légitimées. Ces examens sont donc un outil de normalisation et de légitimation des pratiques des donateurs.

Pour réaliser ces revues, les examinateurs du CAD-OCDE se fondent sur la documentation volontairement fournie par les donateurs. La transparence avec laquelle est transmise cette documentation varie beaucoup selon les pays, en fonction notamment de l'utilité perçue par les responsables politiques d'un tel exercice d'apprentissage et de contrôle mutuels. Les pays d'Europe du Sud ont la réputation d'être moins coopératifs que les pays nordiques et anglo-américains[1245]. On voit ainsi que les donateurs ayant les « meilleures » pratiques sont plus transparents que ceux dont les pratiques sont jugées comme moins satisfaisantes, ces derniers ayant peur d'une mauvaise évaluation qui nuirait à leur réputation internationale. Les examens par les pairs se basent également sur des entretiens auprès du personnel des agences d'aide, au sein des sièges et des agences de terrain, bien que, compte tenu de contraintes de temps et de budget, les visites de terrain se font sur un temps très réduit (quelques jours)[1246].

Les examens par les pairs sont ainsi l'un des terreaux de la diffusion de bonnes pratiques sur l'aide au développement par le CAD-OCDE. À mesure que ces bonnes pratiques se sont institutionnalisées dans le régime de l'aide, les examens par les pairs ont pris un caractère plus normatif, de contrôle, servant à vérifier les progrès faits par les donateurs dans leur mise en œuvre. Le fait que les membres du CAD-OCDE aient ratifié un ensemble de bonnes pratiques permet à l'organisation d'exercer une pression plus importante sur eux. En effet, c'est à partir du moment où les examens par les pairs ont pris cette fonction de contrôle que les donateurs ont commencé à donner de l'importance à leurs jugements et à leurs recommandations[1247]. De plus, leur publication sur Internet en fait un enjeu de réputation internationale pour les donateurs.

Toutefois, ces examens évaluent seulement l'efficacité des donateurs à l'aune des indicateurs globaux fixés par les Hauts Forums sur l'efficacité de l'aide et les OMD, et non par

[1244] Constat fait à partir de l'étude des rapports des examens par les pairs publié par le CAD-OCDE de 2000 à 2013.
[1245] Entretien avec un analyste des politiques au sein du département de l'efficacité de l'aide du CAD-OCDE, 2 juillet 2014.
[1246] VIELAJUS, Martin, *et al*, *Le défi de la « redevabilité » des agences de développement dans leurs propres pays et face à leurs pairs, une étude comparative,* AFD, IRG, GRET, ODI, 2010.
[1247] Cf. Chapitre 3.

rapport à leurs résultats concrets dans les pays bénéficiaires. Le point de vue des bénéficiaires sur l'aide qui leur est apportée est également peu investigué[1248].

Jusqu'au milieu des années 2000, les examens par les pairs du CAD ont été le seul mécanisme international d'évaluation systématique des politiques d'aide des donateurs bilatéraux. Bien qu'elle ait été de plus en plus poussée, la redevabilité internationale qui pesait sur les donateurs se limitait donc au cercle des pairs du CAD-OCDE. En effet, l'aide bilatérale a longtemps été considérée comme une politique publique nationale sur laquelle les États donateurs étaient souverains et n'avaient pas à rendre compte à des acteurs étrangers.

- **Une responsabilité partagée de l'APD pour garantir les biens publics mondiaux**

Cette approche a changé quand l'aide bilatérale a été considérée, non plus comme une politique publique nationale, mais comme une politique servant des biens publics mondiaux relevant d'une responsabilité collective des États du Nord et du Sud.

Les biens publics sont « des biens ou des services qui présentent deux caractéristiques : la non-rivalité - la consommation d'un bien par un individu n'empêche pas sa consommation par un autre - et la non-exclusion - personne n'est exclu de la consommation de ce bien qui est à la disposition de tous »[1249]. À l'échelle internationale, il s'agit de biens très étendus que possède et dont devrait jouir toute l'humanité (par exemple, un accès à l'eau potable, un air sain), et dont la responsabilité est partagée par cette humanité tout entière.

Mais cette responsabilité diffuse a pour conséquence qu'aucun acteur ne souhaite la porter, la déléguant aux autres, comme l'exprime l'ancien secrétaire général des Nations unies Kofi Annan : « une paix plus solide, une prospérité mieux partagée, un environnement épargné : rien de ceci n'est hors de portée si l'on en a la volonté politique. Mais, ni les marchés, ni les gouvernements ne souhaitent se charger de la réalisation de ces objectifs. C'est pourquoi nos efforts doivent se tourner vers le terme manquant de l'équation : les biens publics à l'échelle mondiale »[1250]. Kofi Annan donne alors à l'aide au développement la fonction de garantir l'accès de ces biens publics mondiaux aux plus pauvres par la réduction de la pauvreté, le renforcement des États fragiles, la protection de l'environnement, etc.

[1248] VIELAJUS, 2010, *op. cit.*
[1249] GABBAS, Jean-Jacques, HUGUON, Philippe, « Les biens publics mondiaux et la coopération internationale », *L'Économie politique*, vol 4, n°12, 2011, pp. 19-31.
[1250] Kofi Annan, Secrétaire général des Nations unies, discours à New York le 1 mars 1999.

L'aide au développement étant ainsi devenue une politique dont la responsabilité n'est plus nationale mais internationale, il est donc légitime qu'elle soit soumise à une redevabilité internationale étendue. Cette approche du développement des pays du Sud en termes de biens publics mondiaux a entraîné la création d'objectifs internationaux (les OMD), ainsi que de bonnes pratiques qui ont rendu les États redevables les uns envers les autres. C'est alors que l'évaluation de l'aide bilatérale a dépassé le cadre du CAD-OCDE, d'autres acteurs se saisissant de ce mandat.

En effet, dès lors que l'APD a été appréhendée comme un bien public mondial, le fait que les donateurs bilatéraux soient peu soumis à une redevabilité internationale a été de plus en plus critiqué par la société civile internationale (ONGs, think-tanks, centres de recherche, etc.). D'autant plus que les pays bénéficiaires de l'aide sont aujourd'hui soumis à une forte demande de redevabilité de la part des donateurs. Il a alors été perçu comme légitime que les donateurs se soumettent, eux aussi, à une exigence de transparence, de redevabilité et de suivi-évaluation de leurs politiques d'aide auprès d'un cercle plus large que leurs pairs du CAD-OCDE. Les ONG internationales ont été les premières à dénoncer cette inégalité en termes de transparence de l'aide entre, d'une part des pays bénéficiaires sujets à de récurrentes évaluations de leurs politiques de développement par les donateurs et, d'autre part des donateurs dont les politiques d'aide restent opaques. Lors du Haut Forum sur l'efficacité de l'aide d'Accra en 2008, les ONG ont plaidé pour la création d'une initiative internationale renforçant la transparence de tous les donateurs. C'est ainsi qu'est née l'*International Aid Transparency Initiative* (IATI), une plateforme gérée par l'ONG *Publish what you fund* créée à cette occasion, sur laquelle les donateurs se sont engagés à communiquer leurs données en termes d'allocation d'aide au développement (engagements, prévisions, déboursements, conditionnalités, répartition par pays, répartition sectorielle, etc.)[1251].

C'est dans ce contexte d'attente généralisée d'une meilleure analyse critique des politiques d'aide des donateurs, permettant enfin d'appréhender leur efficacité et leur qualité, mais aussi de comparer l'efficacité des organismes d'aide les uns par rapport aux autres, que quelques chercheurs se sont intéressés à la question de l'évaluation des politiques d'aide des donateurs.

En effet, alors que beaucoup de chercheurs se sont penchés sur la question de l'évaluation de l'efficacité générale de l'aide au développement, les travaux académiques cherchant à

[1251] *International Aid Transparency Initiative Accra Statement*, Ghana, 4 septembre 2008.

évaluer la qualité des politiques des donateurs étaient inexistants jusqu'au milieu des années 2000. Seules les stratégies d'allocation de l'aide des donateurs étaient un objet de recherche récurrent[1252]. En 2006, David Roodman[1253], du *Center for Global Development* (CGDev), a mis en place une méthode d'évaluation de la qualité des politiques de coopération des donateurs à partir de l'indice *Commitment to Development* créé par le CGDev en 2003. Cet indice représente le degré de cohérence entre les diverses composantes de la politique étrangère des États de l'OCDE, à savoir : l'aide au développement, le commerce vers les pays en développement, la finance vers les pays en développement, la politique migratoire, la politique environnementale, le transfert de technologie et la politique sécuritaire. Dans cet indice, la qualité de l'aide au développement est évaluée en termes purement quantitatifs, par le ratio entre le PIB des pays de l'OCDE et le volume d'aide qu'ils allouent aux pays en développement. Dans son évaluation, David Roodman ajoute à ce critère quantitatif plusieurs autres facteurs de qualité de l'aide : la bonne gouvernance et le niveau de pauvreté des pays bénéficiaires, le déliement de l'aide, la réduction de la dette, la taille des projets et des programmes d'aide et la mise en place de mesures fiscales encourageant les dons privés aux pays en développement.

À la suite des premiers travaux du CGDev, William Easterly et Tobias Pfutze[1254] ont élaboré, en 2008, une autre méthode d'évaluation de la qualité de l'aide des donateurs. Ils considèrent quatre catégories de bonnes pratiques : la spécialisation, qui mesure le degré de fragmentation de l'aide parmi les donateurs et parmi les secteurs dans lesquels ils s'investissent ; la sélectivité, qui appréhende l'attachement des donateurs à éviter que leur aide finance des gouvernements autocratiques et corrompus et à ce qu'elle aille dans les pays les plus pauvres ; l'efficacité des canaux d'acheminement de l'aide, qui prend en compte la part d'aide non liée et qui déduit la part de l'aide sous forme de nourriture et d'assistance technique (considérée comme inefficace par les auteurs) ; et la diminution des frais généraux, c'est-à-dire les coûts administratifs de l'aide par rapport au volume d'aide alloué.

Ces deux approches lient étroitement la qualité et la sélectivité de l'aide vers les pays les plus pauvres et les mieux gouvernés. Les chercheurs adoptent l'approche d'une politique d'aide méritocratique qui doit aller vers les PMA les plus volontaires à adopter les « bonnes pratiques » de gouvernance recommandées les donateurs occidentaux. Ces deux méthodes se concentrent encore beaucoup sur l'allocation de l'aide, en délaissant la qualité de la mise en

[1252] *Cf.* Chapitre 8.
[1253] ROODMAN, David, *An Index of Donor Performance*, Working Paper n°67, CGDev, novembre 2006.
[1254] EASTERLY, William, PFUTZE, Tobias, « Where Does the Money Go? Best and Worst Practices in Foreign Aid », *The Journal of Economic Perspectives*, vol. 22, n°2, printemps 2008, pp. 29-52.

œuvre des projets et programmes. À la suite de ces travaux, jugés peu satisfaisants aussi bien par les donateurs que par la société civile, en attente d'outils permettant d'évaluer la qualité et les résultats de l'aide sur le terrain, d'autres travaux de recherche plus complets ont été mis en œuvre.

Stephen Knack, Halsey Rogers et Nicholas Eubank ont notamment tenté, en 2010, d'évaluer de manière « plus compréhensive et représentative les pratiques des donateurs »[1255]. À cet effet, ils se sont concentrés sur quatre « bonnes pratiques » de la Déclaration de Paris : la sélectivité, l'alignement, l'harmonisation et la spécialisation de l'aide au développement, c'est-à-dire le degré de concentration des programmes d'aide sur un groupe de pays bénéficiaires et de secteurs. La qualité de la mise en œuvre des programmes d'aide est donc davantage explorée que dans les travaux de recherche précédents, à travers notamment l'étude des dimensions d'alignement et d'harmonisation. Toutefois, il n'est toujours pas question d'évaluer l'efficacité de la contribution des donateurs au développement des pays bénéficiaires. Dans un contexte de l'aide où la demande de redevabilité est très forte, où il y a une attente de preuves de résultats et où la confiance en l'efficacité de l'aide est affaiblie, cette incapacité à réaliser une évaluation indépendante des performances des donateurs est ressentie comme un manque considérable.

En 2008, à partir des travaux de David Roodman, le CGDev et le *Brookings Institute* se sont associés pour élaborer le *Quality of Development Aid Assessmment* (QUODA)[1256], une méthodologie d'évaluation des politiques d'aide plus complète que les précédentes. La démarche est différente, dans la mesure où il s'agit d'un processus continu d'évaluation, l'exercice étant répété tous les quatre ans, afin de montrer l'évolution de la qualité de l'aide.

L'évaluation couvre à la fois l'efficacité organisationnelle des acteurs de l'aide, l'efficacité de la mise en œuvre de l'aide et l'efficacité en termes de résultats de développement. L'évaluation est donc effectivement plus large que dans les travaux précédents et elle semble plus apte à répondre à la demande politique d'évaluation de résultats des donateurs dans les pays bénéficiaires.

Les 31 indicateurs sont divisés en quatre catégories présentées dans le tableau ci-dessous :

[1255] KNACK, Stephen, ROGERS, Halsey, EUBANK, Nicholas, *Aid Quality and Donor Rankings*, Policy Research Paper, The World Bank, Development Research Group, Human Development and Public Service Team, mai 2010.
[1256] Tous les rapports QUODA sont disponibles sur le site Internet du CGDev : http://www.cgdev.org/section/publications?f%5B0%5D=field_document_type%3A2023; consulté le 21 août 2014.

Tableau 5 : Critères d'évaluation QUODA de l'efficacité des donateurs

Maximisation de l'impact	Renforcement des institutions	Réduction du "fardeau"	Transparence et apprentissage
Part de l'aide allouée à des pays pauvres‡	Part de l'aide allant aux priorités de développement des pays partenaires*‡	Importance des relations de l'aide‡	Signataire de l'IATI‡
Part de l'aide allouée à des pays bien gouvernés‡	Évitement des unités parallèles de mise en œuvre de l'aide*†	Fragmentation entre les agences d'aide‡	Mise en œuvre des standards de collecte et de rapports de données d'IATI
Faibles coûts administratifs‡	Part de l'aide enregistrée dans les budgets nationaux des pays partenaires*†	Taille moyenne des projets*‡	Enregistrement des titres et des descriptions des projets
Haute part d'aide programmable au niveau des pays‡	Part de l'aide aux partenaires ayant des bonnes stratégies opérationnelles‡	Contribution aux organisations multilatérales‡	Description détaillée des projets
Spécialisation des activités d'aide par pays bénéficiaires*‡	Utilisation des systèmes des pays bénéficiaires*†	Missions coordonnées*†	Enregistrement et publication des canaux d'allocation de l'aide
Spécialisation des activités d'aide par secteurs*	Part de l'aide promise enregistrée comme ayant été reçue par les pays bénéficiaires*†	Usage de l'aide programmée*†	Qualité de la politique d'évaluation
Soutien à la fourniture de biens publics mondiaux‡	Coordination de la coopération technique*†	Travail analytique coordonné*†	Complétude des données relatives aux projets
Part d'aide déliée*†	Prévisibilité de l'aide*‡		Aide aux partenaires ayant de bons cadres de suivi-évaluation‡

** = données provenant des gouvernements bénéficiaires; † = de la Déclaration de Paris ‡ = de la littérature académique, ( ), inconnu.*

Source : BIRDSALL, Nancy, KHARAS, Homi, The quality of official development assistance, (QUODA), Third Edition, CGDev, Brookings Institute, 2014

Les sources utilisées pour renseigner ces indicateurs sont des données quantitatives et des informations qualitatives fournies par les gouvernements et la société civile des pays bénéficiaires des programmes d'aide, les documents de suivi de la Déclaration de Paris sur l'efficacité de l'aide et la littérature académique sur l'aide au développement. Elles sont donc relativement diversifiées[1257].

L'analyse de ces indicateurs montre qu'en réalité le QUODA consiste davantage en une évaluation du respect des bonnes pratiques pour parvenir à avoir les meilleurs résultats possibles, qu'en une évaluation des effets des activités des organismes d'aide. L'évaluation ne permet donc pas vraiment de mesurer la qualité de l'aide au développement en fonction des résultats qu'elle produit, ni de comparer les résultats entre les organisations d'aide en termes de réduction la pauvreté dans les pays en développement.

[1257] BIRDSALL, Nancy, KHARAS, Homi, *The quality of official development assistance, (QUODA),* Third Edition, CGDev, Brookings Institute, 2014.

Cette difficulté persistante à réunir les données nécessaires à l'évaluation de l'efficacité des organismes d'aide provient du fait que les chercheurs ne peuvent se baser que sur les données et documents fournis volontairement par ces organismes. Or, ces derniers manquent parfois eux-mêmes de données sur l'efficacité générale de leurs activités, en raison de suivis et d'évaluations incomplets qui ne traitent souvent que des processus et des réalisations des programmes d'aide et non de leurs résultats et de leurs impacts[1258].

Les critères d'évaluation retenus reflètent l'image de ce que serait idéalement une aide au développement de qualité pour les chercheurs du CGDev et du *Brookings Institute*. Premièrement, l'aide devrait être allouée aux pays les plus pauvres et les mieux gouvernés. Deuxièmement, l'aide devrait être adaptée au contexte, prévisible, déliée et mise en œuvre de manière coordonnée avec les autres donateurs, mais aussi avec les pays partenaires. Pour cela, les grands programmes, réalisés conjointement par plusieurs donateurs, ou *via* des organisations multilatérales, devraient être privilégiés. Troisièmement, l'aide devrait avoir comme priorité le renforcement des institutions nationales dans une logique de durabilité. Quatrièmement, les donateurs devraient fournir une information quantitative exhaustive sur leurs activités aux autres donateurs et aux pays bénéficiaires. La redevabilité serait donc un facteur clé de la qualité de l'aide, qu'elle ait lieu entre les différents donateurs, ou entre les donateurs et les pays partenaires. L'aide devrait être rapportée, suivie et évaluée par tous les acteurs impliqués, qui devraient ensuite partager entre eux les informations récoltées pour entraîner un effet d'apprentissage mutuel. Nous retrouvons également l'idée d'une aide qui se concentrerait sur un nombre restreint de pays et de thématiques dans une logique de complémentarité entre les donateurs.

Après l'adoption du Partenariat global pour le développement, les chercheurs du CGDev et du *Brookings Institute* ont souhaité ajouter des indicateurs à l'évaluation QUODA, afin d'être plus en phase avec cette nouvelle doctrine de l'efficacité du développement. Les futurs rapports prendront en compte : le soutien des donateurs au secteur privé et à la société civile, leur respect de la « bonne pratique » de redevabilité mutuelle et leur prise en compte du genre. L'initiative QUODA renforcera ainsi sa dimension évaluative, dans la mesure où elle portera davantage sur le contenu des activités d'aide, plutôt que sur la manière dont elles sont gérées.

Ces initiatives d'évaluation provenant du milieu de la recherche sont perçues comme étant plus indépendantes que les examens par les pairs du CAD-OCDE[1259]. Ce sont également

[1258] Ces différents termes sont définis dans le lexique en annexe 7.
[1259] Entretien avec un évaluateur de la Banque mondiale, 21 mai 2014.

les seuls travaux offrant une évaluation de la qualité de l'aide fournie par les organisations multilatérales et par les donateurs bilatéraux. Nous notons, toutefois, qu'aucune de ces évaluations ne donne une appréciation de l'efficacité des agences d'aide en termes de contribution à l'atteinte des objectifs de développement des pays bénéficiaires. Elles se limitent à l'analyse de l'efficacité organisationnelle des donateurs et à l'étude de l'efficacité des processus de mise en œuvre de leurs activités.

B. Une évaluation grandissante de l'efficacité des organisations multilatérales

Les agences d'aide bilatérales atteignent-elles leurs objectifs sur le terrain ? L'aide est-elle globalement efficace ? Il est impossible de répondre à cette question à partir des évaluations disponibles portant sur les politiques des donateurs. Nous retrouvons la difficulté, observée tout au long de notre recherche, d'avoir une vision globale de l'efficacité de l'aide au développement, et ce malgré une multiplication exponentielle des dispositifs d'évaluation à partir du milieu des années 2000. Les donateurs bilatéraux sont conscients de ce manque, qui leur apporte un certain confort quand il concerne leurs propres politiques, les préservant de voir affichées leurs éventuelles mauvaises performances sur la scène internationale sans qu'ils puissent contrôler cette information, ainsi que de devoir rendre des comptes aux pays bénéficiaires et à la « société civile internationale ». Toutefois, ce manque les handicape dans leur devoir de redevabilité nationale lorsqu'il concerne les organisations multilatérales qu'ils financent. C'est pour combler ce manque d'information sur les organisations multilatérales que certains donateurs bilatéraux ont mis en place des évaluations de la qualité de l'aide fournie par ces donateurs.

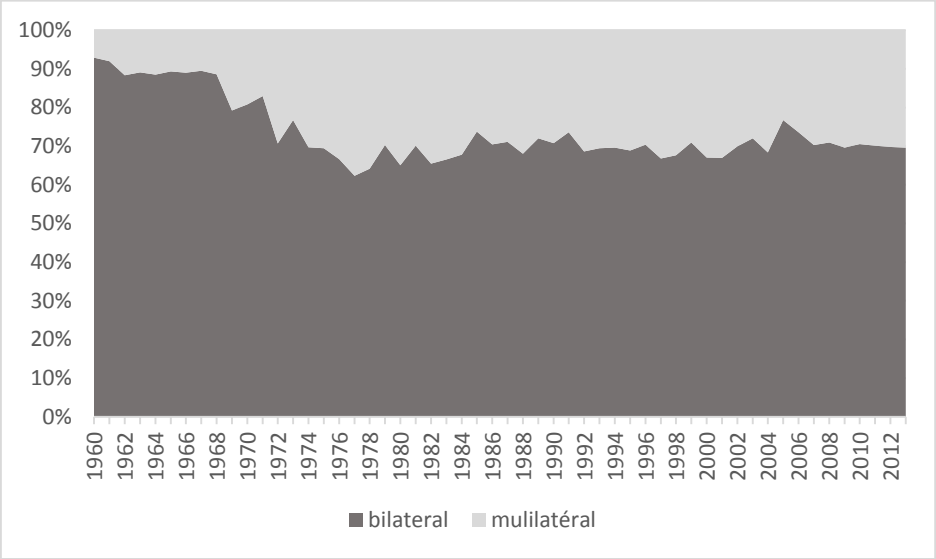
- **Contrôler l'action des organisations multilatérales**

Depuis quelques années, de nombreux dispositifs d'évaluation des politiques des donateurs sont apparus à l'initiative des donateurs bilatéraux. Ils portent tous sur l'évaluation de la qualité de l'aide des organisations multilatérales et non sur celle des autres donateurs bilatéraux, tâche réalisée par les examens par les pairs du CAD-OCDE et par QUODA. Les donateurs bilatéraux déclarent chercher, par ces évaluations, à rationaliser leurs allocations

d'aide aux organisations multilatérales et à améliorer leur redevabilité[1260]. En d'autres termes, ils souhaitent concentrer leurs fonds sur les organisations les plus performantes.

Cet engouement des acteurs de l'aide pour l'évaluation de la qualité de l'aide fournie par les organisations multilatérales s'explique, selon certains, par la multilatéralisation de l'aide au développement, c'est-à-dire par le fait que les donateurs étatiques alloueraient une part de plus en plus grande de leur aide sous une forme multilatérale[1261]. Or, comme le montre le graphique ci-dessous, ce phénomène global de multilatéralisation est peu ma

Graphique 11 : Part de l'aide mondiale bilatérale/multilatérale



Source : OCDE QWIDS, consulté le 11 août 2014

Ce graphique montre que dans les années 1960, l'aide multilatérale représentait environ 10 % de l'aide au développement, alors qu'elle représente aujourd'hui environ 30 %. Toutefois, cette hausse date, pour la plupart des donateurs, du milieu des années 1970 et la tendance est depuis à la stagnation. Il ne nous semble donc pas pertinent de parler d'une multilatéralisation récente de l'aide au développement et donc d'expliquer la hausse des évaluations par les donateurs bilatéraux des donateurs multilatéraux qu'ils financent par ce phénomène.

Nous défendons plutôt la thèse que, dans un contexte où la redevabilité est peut-être la plus importante des bonnes pratiques de l'aide, le fait que les donateurs ne puissent pas rendre compte de l'utilisation et des résultats de 30 % de leurs fonds d'aide pose problème. En effet,

[1260] Entretien avec un ancien agent du DFID travaillant désormais au ministère des affaires étrangères français, 19 juin 2014 ; entretien avec une évaluatrice du DFID, 12 juillet 2013.
[1261] OLIVIER, Guillaume, (dir), *L'aide au développement, un outil à réinventer*, Éditions Charles Léopold Mayer, 2004, 177p, pp. 49-50

au début des années 2000, en l'absence de mécanismes d'évaluation des organisations multilatérales, une large part de l'utilisation des fonds d'aide des donateurs bilatéraux échappait à leur évaluation. Ils n'avaient souvent aucune idée de ce à quoi ces fonds servaient sur le terrain et comment ils contribuaient à l'atteinte des OMD ou des objectifs de leurs politiques d'APD[1262]. En conséquence, l'effort de redevabilité que produisaient les agences d'aide bilatérales, en renforçant leurs pratiques d'évaluation, demeurait incomplet. Les organisations multilatérales réalisaient bien des évaluations, mais celles-ci portaient sur l'ensemble de leurs activités, financées par des fonds propres, privés, de donateurs bilatéraux, de fondations, etc. Cela ne comblait donc pas le souci de redevabilité des donateurs bilatéraux.

C'est pourquoi les donateurs ont décidé de créer des plateformes internationales, afin de se réunir pour évaluer la qualité de l'aide gérée par les organisations multilatérales. Ce phénomène est symptomatique d'un régime de l'aide au développement où les donateurs bilatéraux sont soumis à une pression accrue de résultats et cherchent à concentrer leurs allocations d'aide sur les pays ou les organisations les plus efficaces et rentables. Ainsi, la sélectivité de l'aide au développement ne touche pas seulement les pays du Sud, mais également les organisations multilatérales. Nous observons, en effet, depuis le début des années 2000, un phénomène de concentration de l'aide au développement par donateur sur un nombre de plus en plus restreint de bénéficiaires (pays en développement et organisations multilatérales). Cette concentration est d'ailleurs considérée comme une bonne pratique par le CAD-OCDE dans une optique de rationalisation des politiques d'aide des donateurs, de complémentarité entre les bailleurs de fonds et d'effets d'échelle[1263]. Nous avons montré que, dans le cas des pays bénéficiaires, la concentration s'opère sur les pays considérés comme les mieux gouvernés ou sortant de crise, davantage que sur ceux qui obtiennent les meilleurs résultats dans les évaluations.

Nous estimons que la situation est différente en ce qui concerne les organisations multilatérales et faisons l'hypothèse que les donateurs bilatéraux utilisent bien les évaluations pour orienter leurs fonds d'aide vers les plus performantes. Les donateurs bilatéraux se positionnent alors comme des « financeurs » des organisations multilatérales et font valoir leur droit de regard sur la manière dont leurs financements sont gérés et utilisés. Nous allons observer comment cette performance est appréhendée et évaluée par les donateurs bilatéraux.

[1262] Ce constat a été fait de nombreuses fois lors d'observations participantes et d'entretiens au sein d'organismes chargés de l'aide au développement en 2013 et 2014.
[1263] En rassemblant plus de moyens sur un pays ou une organisation donnés, les donateurs ont plus de chance de parvenir à avoir des résultats massifs à court terme, qu'en dispersant des plus petites sommes entre un grand nombre de bénéficiaires.

Le *Multilatéral Organisation Performance Assessment Network* (MOPAN) a été créé en 2002 par 18 donateurs bilatéraux[1264]. Le choix initial des donateurs de se regrouper en une telle plateforme, plutôt que d'évaluer unilatéralement l'utilisation que font les organisations multilatérales de leurs fonds s'explique avant tout par le coût et la longue durée de telles évaluations : près de 600 000 dollars sur une durée de deux à trois ans[1265]. Le MOPAN produit ainsi un rapport annuel qui évalue le fonctionnement de quatre organisations multilatérales sélectionnées par les donateurs membres.

L'initiative du MOPAN n'apporte, toutefois, pas entière satisfaction aux donateurs la constituant, qui sont par ailleurs tous membres du CAD-OCDE. En effet, les évaluations menées par le MOPAN se concentrent sur l'efficacité organisationnelle des organisations multilatérales et abordent peu leur efficacité en termes de résultats, notamment par rapport aux référentiels des OMD, de la Déclaration de Paris et du Partenariat de Busan. C'est pourquoi, selon des témoignages recueillis au sein du CAD-OCDE, les représentants de l'agence canadienne de développement ont proposé, en 2009, que le réseau Evalnet créé une *task team,* afin d'élaborer un mécanisme d'évaluation des organisations multilatérales complémentaire au MOPAN, qui se concentrerait sur l'évaluation de leur résultats. Les Canadiens ont rallié à leur projet leurs pairs du DFID, de la SADEV (département d'évaluation de l'agence suédoise d'aide au développement), d'USAID (États-Unis), de la Banque mondiale et de l'UNEG (département d'évaluation des Nations unies) pour mettre en place le dispositif[1266]. Nous retrouvons, à l'origine de cette initiative, les donateurs traditionnellement les plus impliqués dans la diffusion d'une culture de l'évaluation dans le secteur de l'aide au développement. Le projet a été validé en 2010 et les premières évaluations réalisées de juin à octobre la même année.

Le problème s'est alors posé de la complémentarité du dispositif d'évaluation du réseau Evalnet du CAD-OCDE avec celui du MOPAN. Nous avons mené une analyse comparative de ces deux dispositifs. Alors que le MOPAN se concentre sur l'efficacité organisationnelle des organisations multilatérales, le CAD-OCDE se donne pour mandat d'évaluer leur « efficacité de développement ». Sa grande valeur ajoutée serait d'apporter, enfin, une réponse à la question de l'efficacité des organisations multilatérales dans les pays bénéficiaires. Pour mener leurs évaluations, les membres du MOPAN utilisent des sources diversifiées : la documentation des

[1264] Allemagne, Australie, Autriche, Belgique, Canada, Corée du Sud, Danemark, Espagne, États-Unis, Finlande, France, Irlande, Luxembourg, Norvège, Pays-Bas, Royaume-Uni, Suède et Suisse.
[1265] OECD-DAC Network on Development Evaluation, *Assessing the Development Effectiveness of Multilateral Organizations: Guidance on the Methodological Approach,* juin 2012.
[1266] Entretien avec une évaluatrice du CAD-OCDE, 2 juillet 2014 ; Entretien avec une Senior Policy Advisor au sein du département de l'efficacité de l'aide du CAD-OCDE, 2 juillet 2014.

agences d'aide (rapports de suivi, évaluations, programmes pays, etc.), des questionnaires envoyés à cinq ou six pays bénéficiaires des organisations multilatérales et des entretiens auprès du personnel des organismes multilatéraux et de leurs bénéficiaires[1267]. Le CAD-OCDE, quant à lui, a choisi une méthodologie plus simple. Il utilise exclusivement les rapports d'évaluation déjà réalisés par les organisations multilatérales comme source pour réaliser ses évaluations[1268]. Ce choix a comme avantage de capitaliser sur les travaux déjà effectués et donc d'économiser du temps, de l'argent et de l'expertise. L'OCDE réalise ce que l'on pourrait appeler une méta-évaluation de l'activité des organismes multilatéraux. Toutefois, le succès de cet exercice tient, dès lors, presque exclusivement à la qualité des évaluations menées par les organisations. Or, cette qualité varie fortement entre les organisations, qui n'ont pas toutes la même culture de l'évaluation et dont les structures et les compétences en la matière n'ont pas la même maturité. Il en ressort que l'évaluation de l'efficacité de l'aide de certaines organisations est parcellaire et peu représentative de la réalité des activités qu'elles mènent.

Par exemple, cela a été le cas de l'Organisation Mondiale de la Santé (OMS) que le réseau Evalnet a tenté d'évaluer en 2013. Les évaluateurs ont conclu que « la principale observation de cette revue est que le nombre limité de rapports d'évaluation disponibles et de qualité suffisante ne permet pas de réunir assez d'informations pour tirer des conclusions sur l'efficacité de l'OMS en termes de développement »[1269]. Ainsi, la principale recommandation faite par les évaluateurs a été d'améliorer la pratique de l'évaluation à l'OMS, aussi bien en termes d'indépendance, que de planification, de transparence, de prise en compte du genre et de l'environnement, etc.[1270].

La qualité médiocre ou variable des rapports d'évaluation pose donc de plus en plus problème dans un système de l'aide où ces rapports commencent à être utilisés à des fins de redevabilité internationale. Pour essayer de pallier cette difficulté, le réseau Evalnet du CAD-OCDE a mis en place en 2010, conjointement avec l'UNEG[1271], un dispositif d'évaluation de la fonction d'évaluation des organisations des Nations unies. Nous avons vu précédemment que les agences d'aide bilatérales sont nombreuses à évaluer leurs systèmes d'évaluation, afin d'en

[1267] Multilateral Organisation Performance Assessment Network, *Annual Report*, 2013.
[1268] OECD-DAC Network on Development Evaluation, juin 2012.
[1269] EVALNET CAD-OCDE, *Development Effectiveness Review of the World Health Organization, 2007-2010*, 27 mars 2013.
[1270] *Ibid*.
[1271] Le rôle du Groupe d'évaluateurs des Nations unies est ici de faciliter l'accès des membres d'Evalnet aux travaux d'évaluation et aux acteurs des Nations unies, ainsi que d'aider à l'analyse des données récoltées et de faciliter ensuite la diffusion des recommandations d'Evalnet au sein des unités d'évaluation des organismes des Nations unies.

améliorer la qualité et l'utilité[1272]. Le système de l'aide est donc pleinement entré aujourd'hui dans une logique de surenchère de l'évaluation.

La différence méthodologique entre les approches du MOPAN et du réseau Evalnet porte également sur l'inclusion des pays bénéficiaires dans les processus d'évaluation. Alors que ces derniers sont impliqués durant toutes les étapes du processus d'évaluation pour le MOPAN, ils ne sont consultés qu'au début et à la fin du processus pour le réseau Evalnet. La démarche du réseau est donc moins inclusive, ce qui permet encore une fois de réduire le temps et les coûts des évaluations, mais amoindrit le caractère participatif de la démarche. L'efficacité des organismes multilatéraux et la qualité de l'aide qu'ils fournissent ne sont donc évaluées qu'à partir de l'idée que s'en font les membres de ces organismes, et non de l'opinion qu'en ont leurs partenaires du Sud.

De cette étude comparative, nous concluons que l'analyse du MOPAN est analytiquement plus riche et qu'elle permet un meilleur dialogue entre les partenaires de développement.

Toutefois, les deux mécanismes d'évaluations apparaissent redondants en bien des points. Par exemple, ils analysent tous deux la qualité de la gestion de l'aide, la gouvernance des organisations multilatérales, l'efficience de l'allocation des fonds, etc. Selon nous, en créant son propre mécanisme d'évaluation de la qualité de l'aide des organisations multilatérales, le réseau Evalnet était bien conscient qu'il y aurait un effet de répétition avec le travail du MOPAN. Sa volonté n'était pas de le compléter et de travailler en collaboration avec cette plateforme, comme présenté dans les rapports publics préparatoires à son instauration[1273], mais bien de fusionner, à court terme, les deux initiatives sous l'égide du CAD-OCDE.

Alors qu'en 2009, durant la phase de préparation de l'initiative, la *task team* du réseau Evalnet soulignait « qu'afin de s'harmoniser avec le MOPAN et de permettre aux deux initiatives d'être complémentaires plutôt que redondantes, les revues sur l'efficacité des organisations multilatérales se feront sur la base des rapports du MOPAN »[1274], elle a rapidement plaidé en faveur du rapprochement des deux initiatives[1275]. Ce travail de plaidoyer a été d'autant plus aisé que le réseau Evalnet et le MOPAN sont constitués des mêmes pays membres, à savoir les donateurs occidentaux. Ainsi, en 2013, le MOPAN a signé un protocole

[1272] *Cf.* Chapitre 4.
[1273] Evalnet, CAD-OCDE, 2012, *op. cit.*
[1274] Site du réseau Evalnet du CAD-OCDE : http://www.oecd.org/dac/evaluation/evaluatingmultilateraleffectiveness.htm, consulté le 22 août 2014.
[1275] MELANSON, James, Development Evaluation Division, Foreign Affairs, Trade and Development Canada, *Comparative Study of MOPAN and Evalnet Approaches to Assessing Multilateral Organizations' Development Effectiveness,* fevrier 2014.

d'accord avec le CAD-OCDE, selon lequel ce dernier assurera désormais le secrétariat du MOPAN. Ce protocole précise que « le MOPAN reste un réseau indépendant et garde la pleine propriété du travail mené ou financé par le MOAPN. Le programme de travail du secrétariat et le budget du MOPAN continuent à être décidés par le MOPAN »[1276].

Ce maintien de deux structures complémentaires et indépendantes l'une de l'autre a, toutefois, rapidement été ébranlé par une évaluation indépendante du MOPAN, réalisée à l'initiative de la *task team* du réseau Evalnet en 2013, dont une des recommandations était la fusion des deux initiatives d'évaluation. Cette recommandation était justifiée dans le rapport par le fait que « le MOPAN reste pertinent et est parvenu à remplir certains des besoins d'information exprimés par ses membres à des coûts raisonnables. Toutefois, l'utilité des évaluations du MOPAN est limitée en termes d'amélioration de l'efficacité des organisations multilatérales »[1277]. Fort de cette conclusion et se proposant de combler les manques du MOPAN, le CAD-OCDE travaille actuellement à la fusion des deux plateformes qui devrait avoir lieu sous peu[1278].

Le CAD-OCDE est ainsi parvenu à englober la plateforme du MOPAN et à renforcer son statut d'organisme de référence sur l'évaluation de l'aide. Il avait déjà dans son mandat l'évaluation de la qualité de l'aide fournie par les donateurs bilatéraux, à laquelle s'ajoute donc aujourd'hui celle des organisations multilatérales. Nous considérons que ce choix de fusionner l'initiative du MOPAN à celle du réseau Evalnet sous l'égide du CAD, au lieu d'élargir le mandat du MOPAN à l'évaluation des résultats des organisations multilatérales, relève plus d'une décision politique de renforcer le positionnement de l'OCDE parmi les organisations internationales, que d'une décision se fondant sur la meilleure qualité des travaux du réseau Evalnet. Un autre argument ayant motivé la décision de privilégier la méthodologie du réseau Evalnet est le gain qu'elle permet en termes de coûts et de durée (125 000 dollars et 110 jours contre 350 000 dollars et 320 jours pour évaluer une organisation[1279]), ainsi que de mobilisation d'expertise. Nous estimons, toutefois, que le portage des évaluations par le CAD-OCDE, organisme dont l'expertise est très reconnue et légitime, aide à une meilleure prise en compte et utilisation des recommandations des évaluations par les organisations multilatérales.

[1276] MOPAN, 2013, *op. cit.*
[1277] BALOGUN, Paul, *Evaluation of MOPAN*, Swedish Institute for Public Administration, août 2013.
[1278] MOPAN, 2013, *op. cit ;* MELANSON, James, 2014, *op. cit.*
[1279] MELANSON, James, 2014, *op. cit.*

Nous avons ensuite analysé plus en détail la méthodologie d'évaluation adoptée par le réseau Evalnet, afin de voir si elle apporte bien une réponse à la question de « l'efficacité de développement » des organisations multilatérales.

Les critères d'évaluation de la qualité de l'aide des organisations multilatérales retenus par le réseau Evalnet sont : l'atteinte des objectifs de développement et des résultats attendus ; l'atteinte d'un développement inclusif et soutenable, grâce à la prise en compte des enjeux transversaux, tels que le genre, l'environnement et la bonne gouvernance ; la durabilité des résultats ; la pertinence des interventions ; l'efficience des activités d'aide et l'utilisation du suivi-évaluation pour améliorer l'efficacité de l'aide. L'approche adoptée est donc réellement différente des autres initiatives d'évaluation. Ce n'est plus seulement l'efficacité de la gestion de l'aide qui est évaluée, mais aussi l'efficacité de l'aide elle-même, c'est-à-dire sa contribution au développement des pays du Sud. La consultation des évaluations réalisées par les évaluateurs du CAD-OCDE montre qu'elles permettent bien d'apprécier si les organisations multilatérales atteignent les objectifs de développement qu'elles se fixent. Ces évaluations sont les premières à fournir un jugement sur l'efficacité des organisations multilatérales à avoir des résultats de développement dans les pays du Sud, bien que ce jugement se fonde peu sur la perception des populations et des gouvernements bénéficiaires.

C'est un grand pas pour l'évaluation de l'aide qui parvient, enfin, à répondre à la demande internationale d'une plus forte redevabilité externe des organismes d'aide sur leurs résultats concrets. Toutefois, cela constitue également une menace pour ces organismes qui voient leurs forces et leurs faiblesses rendues publiques. Étant donné que seulement cinq organisations[1280] ont été évaluées à ce jour par le réseau Evalnet, dont l'OMS pour laquelle l'évaluation n'a pas pu aboutir à des résultats concluants, il n'est pas possible, à ce jour, d'en tirer des conclusions sur l'efficacité générale de l'aide multilatérale.

Malgré ces mécanismes d'évaluation sous l'égide du CAD-OCDE, certains donateurs bilatéraux (la Suède, le Royaume-Uni et l'Australie), ont souhaité développer leurs propres instruments d'évaluation des organisations multilatérales. En effet, les évaluations que nous avons étudiées jusqu'alors apportent une vision générale de l'efficacité de ces organisations, mais ne permettent pas aux donateurs bilatéraux d'identifier comment elles utilisent leurs fonds spécifiquement et en quoi elles contribuent à répondre à leurs priorités particulières.

[1280] Organisation mondiale de la Santé, UNICEF, Banque asiatique de Développement, PNUD, Programme alimentaire mondial.

- **Des évaluations « sur mesure » reflétant une absence de consensus international sur l'efficacité de l'aide**

La Suède est le premier donateur à avoir élaboré un processus d'évaluation des organisations multilatérales en 2008. Selon l'agence d'aide SIDA, « l'objectif de ces évaluations est d'assurer que les fonds d'aide suédois sont utilisés efficacement et qu'ils vont aux organisations les mieux placées pour contribuer au développement. Les organisations sont évaluées sur la base de deux critères : dans quelle mesure leurs activités contribuent aux objectifs de développement de la Suède, et avec quelle efficacité elles utilisent les fonds alloués par la Suède »[1281]. Les évaluations, menées en 2008, couvrent une vingtaine d'organisations multilatérales. Elles montrent une volonté de la Suède de démontrer la pertinence de ces allocations aux organismes multilatéraux par rapport aux objectifs d'aide au développement fixés par les décideurs politiques suédois. Toutefois, dans sa communication, la SIDA met plutôt en avant les idées d'amélioration de la connaissance sur les organisations évaluées et d'adaptation du dialogue et des stratégies de SIDA en fonction de leur fonctionnement[1282]. Le tableau ci-dessous présente les critères d'évaluation utilisés par la SIDA pour évaluer la qualité de l'aide au développement des organisations multilatérales :

[1281] Site Internet du Government Offices of Sweden : http://www.government.se/sb/d/11747/a/122004, consulté le 22 août 2014.
[1282] Site Internet du Government Offices of Sweden : http://www.government.se/sb/d/11747/a/122004, consulté le 22 août 2014.

Tableau 6 : Critères d'évaluation des organisations multilatérales

Cohérence avec les objectifs de la SIDA	Efficacité interne	Efficacité externe
respect et renforcement des Droits de l'Homme	la structure de management joue un rôle stratégique dans l'organisation	l'organisation se concentre sur son mandat principal et les objectifs et priorités de ses secteurs d'intervention
soutien à la démocratie et à la bonne gouvernance	les managers "seniors" dirigent l'organisation de manière efficace	les efforts de l'organisation au niveau des pays mènent à des résultats favorables
soutien à l'égalité entre les femmes et les hommes	la structure organisationnelle est adaptée aux opérations menées	l'organisation est active dans les procédures d'harmonisation de l'aide aux niveaux global et des pays
soutien à l'utilisation durable des ressources naturelles et prise en compte de l'environnement dans les programmes d'aide	la politique des ressources humaines est efficace	l'organisation contribue aux plans d'action nationaux et fixe en fonction ses priorités
soutien à la croissance économique	l'organisation est dotée d'un système de gestion axée sur les résultats opérationnels	
soutien au développement social et à la sécurité	l'information issue du système de gestion axée sur les résultats est activement utilisée	
gestion des conflits	le *reporting* de la gestion axée sur les résultats au niveau des opérations est adéquate	
qualité des opérations humanitaires	l'organisation a un système d'évaluation indépendant et efficient	
	la documentation sur les activités de l'organisation est rendue publique	
	l'organisation a des procédures pour identifier, suivre et traiter les cas de corruption	
	le système d'audit est fiable et efficace	
	le processus de passation de marché de l'organisation est transparent et neutre	

Source : site Internet du Government Offices of Sweden : http://www.government.se/sb/d/11747/a/122004, consulté le 22 août 2014

Dans ces critères, nous retrouvons la culture de l'APD suédoise qui met en avant la dimension égalitaire et sociale du développement (respect des Droits de l'Homme, égalité hommes-femmes, protection de l'environnement, etc.). Nous notons également que l'efficacité de l'aide au développement est évaluée en fonction d'autres facteurs : la sécurité et l'aide humanitaire. C'est donc une approche globale et cohérente du développement qui est privilégiée. Il ne s'agit pas pour la Suède de comparer les organisations multilatérales entre elles et donc d'effectuer un classement des organisations financées. L'outil ne se veut pas normatif et n'entend pas juger de manière objective le fonctionnement et les résultats des organisations.

L'importance donnée à la gestion axée sur les résultats, à la transparence et aux procédures de suivi et d'évaluation est également frappante. La qualité de l'aide semble principalement dépendre de ces facteurs. Encore une fois, la mise en place de dispositifs de suivi-évaluation transparents, à tous les niveaux de la mise en œuvre de l'aide au développement, et dont les informations produites sont diffusées et utilisées collectivement par les donateurs, est considérée comme une bonne pratique fondamentale.

La Suède reprend ainsi les « bonnes pratiques » de la Déclaration de Paris. En premier lieu, la gestion axée sur les résultats, qui figure dans trois indicateurs, mais également l'harmonisation et l'alignement sur les plans d'action nationaux. Toutefois, l'appropriation et la redevabilité mutuelle sont absentes du cadre d'évaluation, ce qui révèle la moindre importance donnée à ces « bonnes pratiques » par la Suède et par les donateurs en général[1283]. La satisfaction des pays bénéficiaires, leur intégration dans les programmes d'aide et leurs relations avec les donateurs n'apparaissent pas comme des facteurs de qualité de l'aide des organisations multilatérales.

L'instrument d'évaluation des organisations multilatérales développé par le DFID, en 2011, est quelque peu différent. Premièrement, le DFID affiche clairement que la décision de mettre en place cet instrument a été motivée par la volonté de mieux contrôler l'utilisation des fonds d'aide au développement alloués[1284]. Ce contrôle est devenu une priorité pour le Royaume-Uni, selon le principe de *value for money* qui fonde depuis plusieurs années la politique d'aide au développement du gouvernement. Appliqué au cadre de l'aide au développement, ce principe signifie « maximiser l'impact de chaque *livre sterling* dépensée sur l'amélioration des vies des populations pauvres »[1285]. Le principe d'efficience est exacerbé ; les résultats des programmes d'aide sont évalués à l'aune de l'investissement monétaire qu'ils ont coûté. Il est attendu du DFID qu'il analyse dans ses travaux d'évaluation le ratio coût/efficacité de tous les programmes d'aide qu'il met en place[1286]. Cette approche de l'efficacité de l'aide est symptomatique d'un contexte de crise budgétaire où les fonds disponibles d'aide au développement se raréfient. Elle implique également, comme nous en avions déjà fait l'hypothèse ci-dessus, une concentration des fonds alloués vers les bénéficiaires les plus rentables, ayant la capacité de parvenir à des résultats élevés à partir d'un investissement

[1283] *Cf.* Chapitre 8.
[1284] MITCHELL, Andrew, Secretary of State for International Development, foreword to the Multilateral Aid Review of DFID, mars 2011.
[1285] DFID, *DFID's Approach to Value for Money (VfM),* juillet 2011.
[1286] Entretien avec une évaluatrice du DFID, 12 juillet 2013.

minimal. Ainsi, seule la perspective d'une amélioration rapide de leur capacité à se développer peut reclasser les bénéficiaires les plus faibles.

La mise en place d'un processus d'évaluation de l'efficacité des organisations internationales a été la mesure phare du secrétaire d'État au développement Andrew Mitchell après sa prise de fonction en 2011. Il lie directement cet exercice au besoin du DFID d'améliorer la *value for money* de l'aide au développement, ainsi que sa redevabilité envers les contribuables britanniques[1287]. L'évaluation est donc réalisée dans un objectif de rationalisation budgétaire de l'aide au développement et implique une mise en concurrence des organisations multilatérales, où les moins bien notées se voient menacées d'une diminution, voire d'un arrêt, des financements volontaires dont elles bénéficient. Le travail d'évaluation des organisations multilatérales réalisé par le DFID est reconnu comme étant de grande qualité par les autres donateurs bilatéraux. Ces derniers l'utilisent comme référence quand ils cherchent de l'information sur l'état de fonctionnement et l'efficacité d'une organisation multilatérale ou quand il cherche à évaluer la pertinence de leur allocation de fonds d'aide dans un secteur donné[1288].

Le tableau ci-dessous présente les indicateurs de qualité de l'aide sur lesquels se fonde l'évaluation du DFID :

[1287] MITCHELL, Andrew, Secretary of State for International Development, foreword to the Multilateral Aid Review of DFID, mars 2011.
[1288] Ce constat a été fait lors d'observations participantes au sein d'organismes d'aide au développement en 2013 et 2014.

Tableau 7 : Critères d'évaluation des organisations multilatérales du DFID

Catégories	Indicateurs
rôle important dans l'atteinte des objectifs du DFID	l'organisation a un rôle important dans l'atteinte des principaux objectifs internationaux de développement ou humanitaires et en fournit des preuves au niveau des pays
	l'organisation a un rôle important dans l'atteinte des principaux objectifs britanniques de développement ou humanitaires, en fournissant des preuves au niveau des pays
prise en compte des enjeux transversaux	l'organisation a une bonne performance dans les États fragiles
	l'organisation soutient l'égalité en termes de genre
	l'organisation s'assure que ses activités protègent l'environnement et le climat
concentration sur les pays les plus pauvres	l'organisation alloue ses ressources aux pays qui en ont le plus besoin ou aux zones nécessitant prioritairement un soutien humanitaire
	l'organisation alloue ses ressources aux pays où elles seront le mieux utilisées
contribution aux résultats	les objectifs de l'organisation s'attachent à atteindre les plus pauvres
	les objectifs de l'organisation s'attachent à entraîner des résultats au niveau des pays
	l'allocation des ressources sert de manière adéquate ces objectifs
	l'organisation contribue aux résultats de développement ou humanitaires des pays
gestion stratégique et de la performance	l'organisation a un mandat clair et une stratégie et une planification opérationnelle pour le réaliser
	l'organe de gouvernance assure la redevabilité du management
	la direction de l'organisation est efficace
	les résultats sont mesurés
	l'organisation a une fonction d'évaluation efficace
	l'organe de gouvernance et le management utilisent les évaluations pour améliorer la prise de décision
	les politiques et pratiques de ressources humaines sont de qualité
conscience des notions de "coûts" et de "bénéfices"	l'organisation encourage et soutient ses partenaires à penser en termes de "value for money"
	l'organisation considère que le retour sur investissement et l'efficience des programmes sont des enjeux importants
	l'organisation fait des économies en matière d'inputs des programmes mis en œuvre
	l'organisation contrôle ses coûts administratifs
comportement partenarial	l'organisation travaille efficacement en partenariat avec les autres acteurs du développement
	l'organisation met en place des politiques de responsabilité sociale donnant la parole aux bénéficiaires
	l'organisation a assez de flexibilité pour permettre une gestion des activités par les bénéficiaires
	l'organisation respecte les principes de l'agenda sur l'efficacité de l'aide

	l'organisation a un rôle de chef de file dans la coordination des donateurs sur le terrain
transparence et redevabilité	l'organisation a une politique globale et ouverte de diffusion de l'information
	l'organisation promeut la transparence et la redevabilité auprès de ses partenaires
	l'organisation publie systématiquement la documentation et les données sur ses projets
	l'organisation est signataire de l'IATI et suit les standards de l'initiative
	les structures de gouvernance incluent une représentation effective des pays partenaires
	les pays partenaires ont le droit d'adresser des plaintes à l'organisation
probabilité d'évolutions positives	l'organe de gouvernance et le management cherchent continuellement l'amélioration
	l'organisation a déjà montré sa capacité à se réformer positivement dans le passé
	les opportunités de réforme sont anticipées

Source : DFID, *Multilateral Aid Review, Ensuring maximum value for money for UK aid through multilateral organizations,* mars 2011

Ces indicateurs montrent que le Royaume-Uni lie fortement la sélectivité et la qualité de l'aide, ce que nous ne retrouvions pas dans le cadre d'évaluation de la SIDA. Une aide de qualité est une aide qui se concentre sur les pays les plus à même d'apporter de *la value for money* aux organisations donatrices. Ces indicateurs reflètent la culture politique d'APD du Royaume-Uni, qui se concentre sur la lutte contre la pauvreté, dans l'esprit des OMD.

Comme la SIDA, le DFID donne une grande place aux critères de transparence, de gestion axée sur les résultats et d'évaluation. La redevabilité est donc considérée comme une des clés de la qualité et de l'efficacité de l'aide. Nous trouvons dans l'évaluation du DFID des références plus directes et exhaustives aux « bonnes pratiques » de l'agenda sur l'efficacité de l'aide que dans celles de la SIDA. En effet, le Royaume-Uni est le membre du CAD-OCDE qui s'est le plus engagé dans la conception et la mise en œuvre de cet agenda[1289]. Enfin, contrairement à la SIDA, l'évaluation du DFID se conclut par un classement des organisations multilatérales qui met en exergue la volonté sous-jacente des décideurs politiques britanniques de les mettre en concurrence[1290].

La démarche d'évaluation des organisations multilatérales adoptée par l'Australie, en 2012, se rapproche de celle du DFID[1291]. Contrairement aux autres donateurs, l'Australie s'est

[1289] DFID, *Multilateral Aid Review, Ensuring maximum value for money for UK aid through multilateral organizations,* mars 2011.
[1290] Voir partie 9.3.
[1291] Australian Government, Australian Aid, *Australian Multilateral Assessment,* mars 2012.

notamment engagée dans cet exercice en raison de la hausse récente de la part multilatérale de son aide au développement (37 % du total de l'aide). De plus, cette décision reflète l'influence forte du DFID sur *l'Australian Aid Agency* (Ausaid) qui reproduit régulièrement ses innovations[1292]. Cette démarche s'intègre également dans le cadre de gestion axée sur les résultats adopté par l'Ausaid en 2011. L'Ausaid évalue 42 organisations multilatérales à partir des deux mêmes critères que le DFID, à savoir l'efficacité organisationnelle des organisations et leur contribution aux objectifs de l'APD australienne. L'évaluation contient un parangonnage des organisations multilatérales, opérant ainsi un classement des organisations. Le tableau ci-dessous décrit les critères d'évaluation adoptés par l'Ausaid :

Tableau 8 : Critères d'évaluation des organisations multilatérales de l'Ausaid

résultats et pertinence	efficacité organisationnelle
l'organisation fournit des preuves de résultat en termes de réduction de la pauvreté et de développement durable en accord avec son mandat	l'organisation possède des documents de planification clairs qui guident la prise de décision et dont la mise en œuvre est suivie
les résultats des activités d'aide sont suivis et diffusés	l'organe de gouvernance fournit des orientations stratégiques claires et, si nécessaire, dirige et suit les réformes et les efforts d'amélioration du management
l'organisation fait l'effort de cibler les groupes les plus pauvres et les plus vulnérables dans les pays et régions où elle intervient	l'organisation a un système global de suivi et d'évaluation des performances des programmes qui sert tous les besoins de redevabilité et d'apprentissage et qui est systématiquement appliqué à travers l'organisation pour informer la prise de décision
le mandat de l'organisation est étroitement aligné avec les objectifs de l'APD australienne	les organes décisionnaires de l'organisation ont réussi à conduire des changements pour améliorer les politiques de ressources humaines et faciliter le recrutement des meilleurs personnels possible
l'organisation a une politique concernant les enjeux transversaux du développement dont la mise en œuvre est considérée comme efficace par les évaluations	le management de l'organisation et l'organe de gouvernance prennent systématiquement en compte les enjeux de *value for money* dans leur prise de décision
l'organisation intervient efficacement dans les États fragiles	la notion de *value for money* est une considération systématique dans la planification et la mise en œuvre des programmes, un facteur dans la prise de décision et dans la résolution des problèmes
l'organisation est impliquée et efficace dans la coordination des efforts de développement entre les donateurs	les enjeux de *value for money* sont des considérations systématiques et importantes dans les accords et les interactions avec les partenaires et les progrès sur ce thème sont étroitement suivis
l'organisation joue un rôle dans l'élaboration de normes et de standards qui sont utilisés et perçus comme utiles par les parties prenantes du développement	l'organisation construit des partenariats efficaces avec des partenaires pertinents
la production de connaissances, les politiques et/ou les innovations de l'organisation sont largement utilisées	l'organisation s'aligne de manière appropriée sur les priorités et les systèmes des pays partenaires, lesquels voient cet alignement comme approprié

[1292] NEGIN, Joël, *The DFID-isation of Ausaid*, 2 juillet 2013: http://devpolicy.org/the-dfid-isation-of-ausaid-20130702-2/, consulté le 22 août 2014.

par de nombreux acteurs du développement et perçus comme utiles	
	l'opinion de chaque partie prenante est considérée comme importante et guide la prise de décision. Les parties prenantes sont satisfaites de leur implication et de leur influence dans la prise de décision
	l'organisation suit pleinement les standards de l'IATI en termes de diffusion de l'information
	l'organisation a une politique d'allocation de ses ressources transparente et basée sur des informations objectives
	des systèmes de gestion financière, d'audit, de gestion des risques et de prévention des fraudes sont en place et sont évalués de manière indépendante comme étant efficaces
	la transparence et la redevabilité sont toujours des considérations importantes dans les accords et interactions avec les partenaires de l'organisation

Source : Australian Government, Ausaid, *Australian Multilateral Assessment,* mars 2012

Ces indicateurs reflètent clairement l'influence du DFID sur l'approche de l'évaluation de l'Ausaid. Les références au principe de *value for money* sont très nombreuses, et nous retrouvons une approche de l'aide qui se concentre sur la réduction de la pauvreté. Toutefois, à la différence du DFID, l'Ausaid ne fait pas référence à une sélectivité de l'aide en termes de bonne gouvernance. Le volet évaluant l'efficacité des organisations en matière de résultats de développement est beaucoup plus faible que celui évaluant leur efficacité organisationnelle. Comme dans les autres évaluations, les bonnes pratiques liées à la transparence et au suivi-évaluation sont largement mises en avant. En effet, la diffusion de l'information est la première condition de faisabilité de ce type d'évaluation. La première amélioration demandée aux organisations est donc, logiquement, de renforcer leurs pratiques de collecte de données, de suivi, d'évaluation et de diffusion des informations recueillies par ces différents biais aux autres donateurs.

Nous voyons ainsi que les évaluations des organisations multilatérales menées par les donateurs bilatéraux servent majoritairement un objectif de redevabilité. En effet, en cherchant à prouver que les organisations contribuant le mieux aux objectifs des donateurs bilatéraux et étant les plus efficaces sont les plus financées, les évaluations concluent indirectement que les fonds d'aide sont attribués de manière efficiente et utile. Si ce n'est pas le cas, ces évaluations deviennent un outil de justification des réallocations budgétaires des donateurs bilatéraux vers certaines organisations multilatérales (les plus performantes) au détriment d'autres. Elles

donnent une image neutre et objective à ces décisions. Elles répondent également à une attente généralisée des donateurs bilatéraux de regagner un certain contrôle sur la manière dont les organisations multilatérales utilisent les fonds d'aide au développement qui leur sont alloués[1293].

En plus de ces initiatives d'évaluation par les donateurs bilatéraux, des banques de développement[1294] se sont regroupées en une plateforme d'évaluation par les pairs. Cette plateforme, le *Common Performance Assistance System* (COMPAS), a été créée en 2005. Elle a vocation à évaluer « les capacités des banques multilatérales de développement à suivre un cadre de gestion axée sur les résultats »[1295]. L'apprentissage mutuel est présenté comme le principal bénéfice de ce dispositif. Il s'agit d'identifier et d'échanger les bonnes pratiques que les banques mettent en place pour améliorer la gestion de leurs fonds, ainsi que leurs résultats. Le COMPAS n'est pas conçu pour établir une comparaison directe entre les banques de développement.

Le tableau ci-dessous synthétise les critères inclus dans les évaluations étudiées ci-dessus. Il illustre ainsi les champs recouverts et la vision de la qualité de l'aide reflétée par ces travaux.

[1293] Constat notamment fait pour la France lors d'une évaluation sur l'APD de la France en matière de santé dont la part multilatérale est élevée. Les décideurs politiques français regrettaient que les contributions aux organisations multilatérales de la France soient trop peu suivies et évaluées et qu'elles ne soient pas davantage orientées vers les organisations les plus efficaces.

[1294] Banque africaine de développement, Banque asiatique de développement, Fond international pour le développement agricole, Banque interaméricaine de développement, Banque islamique de développement, Banque européenne pour la reconstruction et le développement, Banque mondiale.

[1295] Site Internet de la plateforme COMPAS : http://www.mfdr.org/Compas/index.html, consulté le 22 août 2014.

Tableau 9 : Synthèse des critères d'évaluation des organisations multilatérales

	CAD Peer Reviews	Quoda	MOPAN	Evalnet-CIDA	SADEV	DFID	AUSAID	COMPAS	Récurrence
volume d'aide, répartition, modalité	oui	oui	oui	oui	oui	oui	oui	oui	**8**
Transparence/suvi-évaluation et redevabilité	oui	oui	oui	oui	oui	oui	oui	oui	**8**
Gestion axée sur les résultats	oui	oui	oui		oui	oui	oui	oui	**7**
alignement	oui	oui	oui		oui	oui	oui		**6**
harmonisation	oui	oui	oui		oui	oui	oui		**6**
Efficacité organisationnelle	oui		oui		oui	oui	oui	partielle ment	**5,5**
Pertinence des objectifs et du mandat de l'organisation			oui	oui	oui	oui	oui		**5**
Prise en compte des enjeux transversaux	oui			oui	oui	oui	oui		**5**
Partenariat avec les pays du Sud	oui	oui	oui			oui	oui		**5**
appropriation	oui		oui			oui	oui		**4**
prévisibilité	oui	oui	oui			oui			**4**
Efficacité externes (résultats de développement)		partielle ment		oui	oui	oui			**3,5**
Value for Money/Efficience	partielle ment			oui		oui	oui		**3,5**
Sélectivité de l'allocation de l'aide	partielle ment	oui				oui	oui		**3,5**
Redevabilité mutuelle	oui					partielle ment	partielle ment		**2**

Source : Auteure

Ce tableau montre clairement que les évaluations prennent en compte, de manière récurrente, certains facteurs de qualité et d'efficacité de l'aide, alors que d'autres sont délaissés. Les évaluations se concentrent principalement sur trois critères : l'analyse quantitative de l'aide, le niveau de transparence et de redevabilité des bailleurs et la mise en place de systèmes de gestion axée sur les résultats. Nous trouvons plusieurs explications à ce phénomène. Tout d'abord, l'enjeu de la hausse des volumes d'aide est toujours très présent dans le système de

l'aide, comme en témoignent les appels récurrents des organisations multilatérales à ce que les donateurs bilatéraux atteignent l'objectif des 0,7 % de PNB consacré à l'aide au développement. En outre, il est plus aisé de mener une analyse quantitative de l'aide au développement qu'une analyse qualitative de ses modalités de mise en œuvre et de ses résultats. De même, améliorer la transparence et mettre en place des structures de gestion par les résultats sont des réformes relativement faciles à mener pour les organisations et à suivre pour les évaluateurs, notamment grâce aux bases de données créées sur Internet. De plus, la pensée actuelle sur l'aide au développement, marquée par l'idéologie du New Public Management, donne une importance centrale à la notion de redevabilité, qui implique des procédures multiples de suivi, de contrôle, d'évaluation et de transparence dans les organisations d'aide au développement.

À l'inverse, d'autres critères sont mis de côté par les évaluateurs, notamment la redevabilité mutuelle, qui n'est réellement prise en compte que dans une procédure d'évaluation. Cela montre que les donateurs se sont peu approprié ce principe. Un autre critère peu évalué est l'efficacité en termes de résultats des activités d'aide, c'est-à-dire les améliorations concrètes qu'elles apportent dans les pays du Sud. Pourtant, ce critère attire de plus en plus l'intérêt des donateurs, c'est même la raison pour laquelle la plupart de ces dispositifs d'évaluation ont été créés. Mais, compte tenu du manque de données disponibles à partir des évaluations internes aux organismes multilatéraux, il reste peu investigué. Le critère de *value for money* est également peu repris dans les évaluations. Toutefois, c'est un concept qui monte en puissance dans le secteur de l'aide au développement, notamment dans les agences bilatérales anglo-américaines et à la Banque mondiale.

La qualité de l'aide en matière d'implication des bénéficiaires est surtout prise en compte à travers le principe d'alignement, qui ne recouvre, cependant, que la cohérence des activités d'aide avec les plans de développement nationaux des pays bénéficiaires ou les documents cadres équivalents. Ces documents étant élaborés sous la tutelle des donateurs et pour les satisfaire, ils sont donc, *a priori,* conçus en adéquation avec les attentes des donateurs et peuvent être en décalage avec les besoins de la population bénéficiaire ou les perspectives politiques réelles des gouvernements. Le principe d'appropriation recouvrirait une dimension plus large de l'implication des bénéficiaires. L'appropriation est atteinte si les bénéficiaires considèrent utiles les programmes de développement introduits par les donateurs et s'ils sont capables de les gérer durablement, une fois les donateurs partis. L'appropriation serait donc plus révélatrice de la qualité de l'aide, telle qu'elle est perçue par les gouvernements et les populations

partenaires, que ne l'est l'alignement. Toutefois, complexe à appréhender, elle est plus rarement analysée et de manière moins approfondie.

C. Un instrument technique aux effets politiques

Nous avons vu que les procédures d'évaluation de la qualité de l'aide des donateurs se multiplient et portent sur un nombre croissant de critères, allant du volume et de la répartition des fonds d'aide au développement, aux résultats des projets et programmes mis en œuvre, en passant par l'efficacité organisationnelle des structures d'aide. Dans un contexte où l'aide est légitimée par les résultats quantifiables qu'elle atteint dans les pays du Sud et où les bonnes pratiques de gestion de l'aide prolifèrent, les donateurs sont soumis à une redevabilité internationale grandissante par leurs pairs et par « la société civile internationale ». Alors que jusque dans les années 1990, cette redevabilité, très limitée, s'effectuait dans un esprit d'apprentissage mutuel, elle est aujourd'hui mise en œuvre pour répondre au besoin des donateurs de fournir des preuves de leur efficacité organisationnelle et sur le terrain de l'aide. Toutefois, ces évaluations ont-elles un effet sur les pratiques des organisations évaluées ? Les classements effectués par ces évaluations ont-ils l'effet escompté de les inciter à atteindre le haut du tableau et donc à mettre en place des réformes parfois substantielles transformant leur culture organisationnelle ? Cette dernière partie sera consacrée à répondre à ces interrogations.

- **Le reclassement des fonds et banques de développement en défaveur des organisations onusiennes**

La première question à laquelle nous allons essayer de répondre est de savoir s'il existe une réelle « pression par les pairs » entre les organisations d'aide au développement, qui donnerait aux évaluations un pouvoir de changement sur les politiques et les pratiques des donateurs. Cela semble être le cas pour le CAD-OCDE, dont les recommandations des examens par les pairs sont prises en compte avec sérieux par les donateurs bilatéraux[1296]. Cette pression par les pairs ne s'est pas construite par la mise en place de mesures obligeant les donateurs à appliquer les recommandations des examens. Il s'agit d'un processus plus diffus. Les donateurs se réunissent depuis plusieurs décennies au sein du CAD-OCDE, où ils échangent leurs

[1296] *Cf.* Chapitres 2 et 3.

expériences et leurs connaissances sur l'aide au développement. Des chercheurs se sont intéressés à analyser cette pression par les pairs au sein du CAD-OCDE : « l'approche institutionnalisée et formelle du CAD, soutenue par une bureaucratie experte, fut la fondation d'une communauté épistémique. Même quand la mise en place des recommandations proposées par les examens par les pairs est impossible, il semble qu'il y ait une certaine pression qui fait qu'il vaut la peine pour le pays évalué de considérer sérieusement d'adapter ses programmes d'aide dans le sens suggéré par les examens »[1297]. Nous avons vu que cette idée se vérifie sans le cas de la Corée du Sud. Aujourd'hui, l'existence de ces examens est bien connue du personnel des agences d'aide, même lorsqu'ils méconnaissent les « bonnes pratiques » de l'efficacité de l'aide produites par le CAD-OCDE, pour qui il est important d'avoir des examens élogieux. De ce fait, les pratiques et les politiques de l'aide au développement ont tendance à s'harmoniser entre les donateurs[1298].

En outre, depuis les années 2000, le CAD-OCDE a institué des bonnes pratiques par des accords internationaux qui engagent les donateurs, ce qui augmente la pression par les pairs et donne aux examens par les pairs un aspect plus contraignant. Bien que certains donateurs s'opposent souvent aux innovations ou aux bonnes pratiques proposées par les donateurs chefs de file du Comité (Royaume-Uni, pays scandinaves, Danemark), ils finissent par suivre la tendance et par aligner leurs discours et, au moins en apparence, leurs pratiques. Cela a été, par exemple, le cas de la France qui s'est d'abord opposée à l'élaboration de la Déclaration de Paris pour finalement faire de ce document un pilier normatif de sa politique d'aide au développement[1299].

Toutefois, le phénomène de l'évaluation de la qualité de l'aide dépasse aujourd'hui le cadre institutionnalisé du CAD-OCDE. Nous pouvons donc nous demander si les évaluations menées par des chercheurs (CGDev-Brookings Institute, William Easterly, Stephen Knack *et al*.) ou des donateurs bilatéraux (Royaume-Uni, Suède, Australie) ont un effet sur les pratiques des donateurs évalués. En effet, ces évaluations ne s'inscrivent pas dans un cadre institutionnel comme les examens par les pairs du CAD-OCDE (excepté la nouvelle initiative du réseau Evalnet). Ainsi, pour renforcer leur influence, ces évaluations sont souvent accompagnées d'un classement des donateurs dans l'espoir que les « mauvais élèves » pointés du doigt soient plus fortement incités à faire évoluer leurs pratiques dans le « bon » sens, c'est-à-dire vers plus

[1297] BEN-ARTZI, Ruth, *Ios and Peer Pressure : An Examination of the Development Assistance Committee (DAC)*, Prepared for 8th Pan European Conference on International Relations, Warsaw, 21 septembre 2013.
[1298] *Cf.* Chapitre 3.
[1299] Entretien avec une évaluatrice du CAD-OCDE, 2 juillet 2014.

d'efficience. Selon Stephen Knack, « ces classements publics ont vocation à inciter les donateurs bilatéraux et multilatéraux à renforcer leurs efforts d'efficacité et d'efficience, afin d'améliorer leur rang »[1300].

Nous avons comparé les différents classements produits par les évaluations de la qualité de l'aide des donateurs. L'objectif était de voir si ces classements convergeaient ou divergeaient. Il faut noter que les donateurs évalués ne sont pas les mêmes dans chaque classement, certains couvrant à la fois les donateurs bilatéraux et multilatéraux, alors que d'autres se concentrent sur les donateurs multilatéraux.

Un aspect important à souligner est que, dans une même évaluation, les classements entre chaque critère sont peu corrélés, c'est-à-dire que des donateurs qui sont très bons sur un critère d'évaluation, par exemple la transparence et le suivi-évaluation, peuvent être médiocres sur un autre critère, tel que la *value for money* ou l'efficacité organisationnelle. Le coefficient de corrélation obtenu entre chaque critère se situe généralement entre 0,2 et 0,3. Ce constat est d'autant plus vrai pour les donateurs classés dans la moyenne, qui ont les résultats les plus variables. Toutefois, la comparaison permise par le tableau ci-dessous apporte des éléments de réflexion intéressants.

[1300] KNACK *et al*. 2010, *op. cit.* p. 2.

Tableau 10 : Classements des donateurs selon leur efficacité

	QUODA	DFID	AUSAID	EASTERLY	KNACK
1	Irlande	GAVI	PAM	IDA	Banque. Asiat. Dev
2	IDA	PIDG	ICRC	Banque. Afr. Dev	Hongrie
3	Banque africaine dev	CICR	Banque mondiale	Banque. Asi. Dev	Pologne
4	Danemark	Fonds Asiat. Dev	UNHCR	IDB	FMI
5	Banque de dev. Int.	EDF	GAVI	Norvège	Banque Africaine. Dev
6	FMSTP	FMSTP	UNICEF	Suède	IDB
7	Royaume-Uni	ECHO	FMSTP	Japon	Banque mondiale
8	Canada	IDA	Fonds. Asia.Dev	Suisse	Australe
9	Suède	UNICEF	UNRWA	Portugal	IFAD
10	Banque asiatique de dev.	Fonds Dev.Afr	OMS	France	CE
11	FIDA	FIDA	Global. Partn. Educ	Australie	GAVI
12	Finlande	CIFs GFDRR	IFC	UNICEF	France
13	Nouvelle Zélande	UNHCR	UNOCHA	Belgique	Suède
14	CE	Glob.Economic Fund	Fonds asiat. Dev	Italie	Danemark
15	Portugal	EBRD	PBF	États-Unis	Royaume-Uni
16	Japon	WFP	UNMAS	Autriche	FMSTP
17	Pays-Bas	Banque asiat de dev	CIFs	Irlande	États-Unis
18	Norvège	PNUD	PNUD	Fonds. Nordique de dev.	Pays-Bas
19	France	CERF	MLF	Pays-Bas	Portugal
20	Grèce	UNITAID	PIDG	Canada	Italie
21	Espagne	OCHA	GEF	Danemark	Canada
22	Australie	IFRC	CGIAR	Finlande	Allemagne
23	Allemagne	FNUAP	GCDT	Luxembourg	Irlande
24	Luxembourg	EFW	FNUAP	FMI	Japon
25	États-Unis	IAD	IFAD	Allemagne	Norvège
26	Suisses	CE	GFDRR	Banque Caraïbe. Dev	Espagne
27	ONU	UNAIDS	OIT	CE	Nouvelle Zélande
28	Italie	OMS	UNAIDS	EBRD	Autriche
29	Belgique	UNEP	PAM	Grèce	Belgique
30	Autriche	Banq. Dev. Caraïbes	UNCDF	PNUD	Finlande
31	Corée du Sud	UNIFEM	FMI	Espagne	ONU
32		IOM	IOM	Nouvelle Zélande	Suisse
33		ONUDI	OHCHR	FNUAP	Luxembourg
34		PAM	UNODC	IFAD	EBRD
35		OIT	IDB	PAM	Grèce
36		UN habitat	LDCF	GEF	Corée du Sud
37		UNESCO	UN-Habitat	UNHCR	Turquie
38		UNISDR	UNISDR		République Tchèque
39		Commonwealth Sec	AF		
40			UNESCO		
41			PNUE		
42			Commonwealth Sec		

Source : auteure

Dans ce tableau, les cases gris clair représentent les donateurs qui se trouvent plusieurs fois dans les 15 premiers rangs et les gris foncé ceux qui sont plusieurs fois dans les 15 derniers. Nous constatons que les banques de développement (Banque mondiale, Banque asiatique de Développement, Banque africaine de Développement), ainsi que les fonds verticaux (le Fonds mondial de lutte contre le sida, la tuberculose et le paludisme (FMSTP), l'Alliance globale pour les vaccins et l'immunisation (GAVI), les fonds asiatique et africain pour le développement) se retrouvent régulièrement en tête des classements. Cela tend à montrer que les critères d'évaluation retenus favorisent les institutions qui fonctionnent selon une logique financière, où la maîtrise des risques, le ratio coûts/résultats et la notion de *value for money* sont très présents dans la culture organisationnelle.

Nous allons démontrer, en prenant l'exemple du secteur de la santé, que cette efficacité supérieure des banques de développement et des fonds verticaux est toute relative. Par exemple, les fonds verticaux en matière de santé sont des outils efficaces sur le court terme pour enrayer une pandémie incontrôlée. Toutefois, sur le long terme, ils tendent à fragiliser les systèmes de santé nationaux en mettant en place, dans une logique de verticalité, des systèmes parallèles d'acheminement et de distribution des médicaments et des vaccins. Ce sont des outils très spécialisés qui se concentrent sur un type d'activité (par exemple, la vaccination pour GAVI) ou sur un enjeu (les pandémies du sida, de la tuberculose et du paludisme pour le FMSTP). Ils conduisent donc à compartimenter la santé, au lieu d'entretenir une approche globale, considérant que tous les enjeux de santé sont interdépendants. En suivant cette logique, les fonds verticaux peuvent, en outre, créer des inégalités en priorisant certains malades par rapport à d'autres. Un malade du sida se fera, en effet, plus facilement traiter qu'un malade du cancer, car les fonds verticaux se focalisent sur la première maladie sans développer de capacités transversales de traitement. Si ces fonds verticaux ont permis de traiter plusieurs millions de malades (il est estimé que le FMSTP a permis, depuis sa création en 2002, l'attribution de traitements anti rétroviraux à 6,1 millions de personnes, le traitement de 11,2 millions de cas de tuberculose et la distribution de 360 millions de moustiquaires imprégnées[1301]), la durabilité de ces résultats est questionnable, dans la mesure où la grande majorité pays bénéficiaires n'ont pas développé la capacité de maintenir ces traitements une fois que les fonds verticaux se seront désengagés.

[1301] Ces résultats sont présentés par le FMSTP sur son site Internet : http://www.theglobalfund.org/fr/about/results/, consulté le 22 août 2014.

En Côte d'Ivoire, par exemple, ces effets négatifs de la montée en puissance des fonds verticaux sur le système national de santé sont évidents : le ministère de la santé et de la lutte contre le sida, sans organigramme depuis plus d'un an, abrite 24 programmes sanitaires cloisonnés selon les sources de financement nationales et internationales. Dans ce pays, la gratuité des soins vaut pour les personnes séropositives, mais pas pour celles qui souffrent de l'hépatite B ou du cancer ; elle reste souvent théorique pour la santé maternelle et infantile. De plus, les hôpitaux et dispensaires de province n'ont pas de personnels et ne fonctionnent plus, car il n'y a plus d'aide à la formation de personnels de santé[1302].

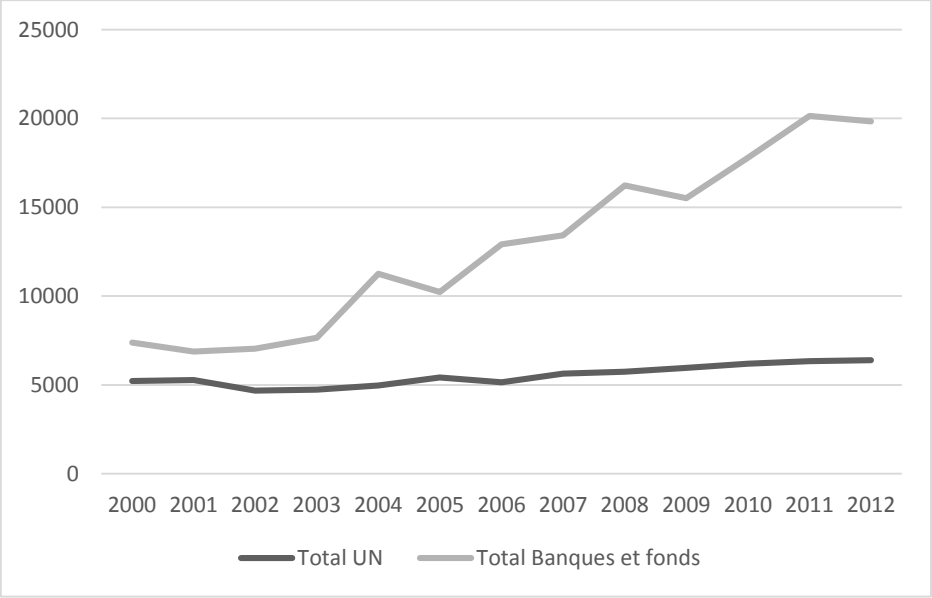
En fin de classement des évaluations, nous trouvons de manière récurrente les organisations des Nations unies, notamment l'UNESCO, UN-Habitat, l'organisation internationale du travail (OIT), etc., ainsi que des petits donateurs bilatéraux, tels que la Corée du Sud, la Nouvelle-Zélande, ou la Grèce. Les Nations unies semblent donc défavorisées par rapport aux institutions financières. Elles sont vues comme étant peu coordonnées, souffrant d'une pesanteur administrative forte les empêchant d'être réactives face aux enjeux émergents du développement mondial et ayant des résultats moins concrets que les banques et fonds de développement. En effet, une partie du mandat des organisations des Nations unies est de produire de l'expertise, de faire du plaidoyer, ou encore de renforcer les capacités institutionnelles des pays en développement, activités qui n'entraînent pas de résultats visibles et quantifiables à court terme. En conséquence, ces organisations sont considérées comme moins utiles. Pourtant, malgré leurs difficultés de fonctionnement internes, les organisations des Nations unies restent des acteurs incontournables de l'aide au développement, offrant un soutien unique aux PMA. Elles interviennent parfois dans des secteurs très spécialisés qui sont délaissés par les autres donateurs, tels que le droit du travail, les migrations ou la culture et la science. Dans le domaine de la santé, l'OMS a un mandat sans pareil de veille sanitaire, de recherche médicale et de production de normes sanitaires qu'aucun autre acteur de développement n'assure. La baisse des financements au profit de l'OMS, compte tenu de son image de donateur peu rentable et peu efficace, entraîne une difficulté plus grande pour l'organisation à remplir son mandat de veille sanitaire sur les grandes épidémies, comme le développement d'Ébola en 2014, ce qui peut mener à des risques sanitaires importants.

Les évaluations reflètent donc bien une conception particulière de l'aide au développement, selon laquelle l'efficacité à court terme et quantifiable prévaut sur l'efficacité à long terme, se fondant sur le renforcement des capacités des pays en développement et leur

[1302] Ces observations ont pu être faites durant une mission de terrain en Côte d'Ivoire en mai 2014.

permettant de devenir autonomes. Nous allons montrer que le déclassement des organisations des Nations unies au profit des banques et fonds de développement contribue à une réallocation de l'aide vers ces derniers. Le graphique ci-dessous compare l'évolution du volume d'aide alloué aux Nations unies et aux banques et fonds de développement :

Graphique 12 : Évolution des contributions aux Nations unies et aux banques et fonds de développement (MUSD)



Source : QWIDS DAC-OECD

Alors que les financements attribués aux banques et fonds verticaux de développement par les donateurs du CAD-OCDE augmentent fortement depuis le début des années 2000, les fonds alloués aux organisations des Nations unies stagnent. Cela laisse penser que si une large partie des contributions aux organisations Nations unies n'était pas obligatoire, elles connaîtraient une baisse marquée. Cette hausse des fonds alloués aux banques de développement et aux fonds verticaux provient, en partie, de la multiplication de ces derniers depuis le début des années 2000. Toutefois, la plus ou moins bonne réputation des organisations multilatérales véhiculée par les évaluations par les pairs et des organismes de recherche a une influence certaine sur la manière dont les donateurs répartissent leurs fonds d'aide au développement. En effet, les classements produits par l'initiative QUODA, ou même ceux réalisés par le DFID ou la Suède, sont largement consultés par les donateurs, qui les utilisent, parmi d'autres instruments, pour décider des montants de leurs financements aux organismes

multilatéraux[1303]. Cet effet de déclassement est plus faible pour les donateurs bilatéraux, dans la mesure où ils sont financés sur les fonds publics de leur État d'appartenance et sont donc moins soumis à la concurrence des autres donateurs.

- **Des évaluations contestées ayant toutefois un effet réformateur**

Un élément d'appréciation de l'influence des classements des évaluations par les pairs est les réactions qu'ils entraînent. À la suite de la publication de la *Multilateral Review Assessment* du DFID, l'UNESCO a rendu publique une lettre condamnant le jugement sévère établi à son égard. Dans cette lettre, elle met en avant le fait que « la revue ne prend pas en compte le mandat particulier des agences spécialisées des Nations unies et leur utilité générale pour la coopération multilatérale. Le score médiocre attribué à l'UNESCO, à partir d'un nombre limité d'indicateurs pertinents par rapport à sa mission, reflète l'échec de tels instruments d'évaluation à capturer les aspects les plus importants du travail de l'UNESCO »[1304]. La lettre précise que « la revue élaborée en 2010 coïncide avec le lancement d'un processus de réformes par la nouvelle directrice générale de l'organisation. Certaines des conclusions sont en accord avec sa propre analyse des faiblesses et manquements de l'organisation, ce qui devrait renforcer le soutien qu'elle recherche de la part du personnel pour mettre en œuvre les réformes nécessaires »[1305]. La réaction de l'UNESCO révèle également la crainte d'un effet d'éviction des financements : « toutefois, l'organisation croit que l'impact négatif de ce classement, à partir d'un ensemble inapproprié d'indicateurs, ne servira pas les intérêts des contribuables britanniques ni du développement international »[1306]. De manière plus générale, l'UNESCO critique le biais de l'évaluation en faveur des fonds verticaux : « la méthodologie utilisée entraîne un biais systématique en faveur des canaux de financement et en défaveur des agences spécialisées des Nations unies, dont les mandats sont plus diversifiés et complexes (que ceux des fonds verticaux), incluant le conseil politique, le renforcement des capacités, la production de normes et de standards, le parangonnage et les opérations de terrain »[1307].

Le Programme des Nations unies pour l'environnement (PNUE), également déclassé par le DFID, a condamné de manière similaire les manquements de l'évaluation qui « reflète une

[1303] Constats issus de documents confidentiels d'organismes d'aide au développement que nous avons pu consulter dans le cadre d'une évaluation à laquelle nous avons participé pour le ministère des affaires étrangères portant sur la coopération française en matière de santé.
[1304] *UNESCO Response to the UK Multilateral Aid Review*, 2011.
[1305] *Ibid.*
[1306] *Ibid.*
[1307] *Ibid.*

vision particulière des objectifs et des paramètres opérationnels de l'aide, qui ne capture pas toujours la raison d'être d'une entité comme le PNUE ayant une fonction normative, ainsi que le mandat de traiter les aspects transversaux de l'environnement et du développement durable »[1308]. Ensuite, le PNUE présente ses principaux résultats depuis 2007, mettant ainsi en avant son efficacité.

De même, l'OIT s'est élevée contre les conclusions de la *Multilateral Review Assessment* du DFID : « le cadre analytique de la *Multilateral Review Assessment* exclut *a priori* toute évaluation positive des organisations multilatérales qui ont un mandat politique et de production de normes, comme c'est le cas de l'OIT. Les critères utilisés par le DFID ne reflètent que les résultats immédiats et tangibles, tels que la construction de routes, l'apport de nourriture aux populations et le nombre de vaccins délivrés. L'accent mis sur ce type de résultats est légitime. Cependant, l'inévitable conséquence est qu'une organisation comme l'OIT, dont le mandat se concentre sur des résultats de développement de long terme, la production de normes et l'appui à la mise en place de politiques publiques, sera toujours évaluée comme inefficace, peu importe la qualité de son travail »[1309]. La réponse de l'OIT prend la forme d'un rapport de 58 pages, consacré à la défense de son utilité et de la qualité de l'aide qu'elle apporte aux pays.

Les organisations des Nations unies ressentent donc un sentiment d'injustice, dans un contexte où l'efficacité de l'aide est évaluée à partir d'indicateurs de résultats de court terme et où les actions de développement social et politique sont délégitimées par rapport aux activités d'urgence, qui relèvent plus de l'humanitaire et à celles portant sur la construction d'infrastructures ou de relance rapide de la croissance économique. En effet, les évaluations ont pris le parti de se focaliser sur les aspects procéduraux de l'efficacité, tels que la gestion axée sur les résultats ou la mise en place de systèmes de suivi-évaluation sophistiqués. Elles prennent la forme d'une analyse majoritairement quantitative des résultats, à partir, par exemple, du nombre de vies sauvées ou du nombre de personnes sorties de la pauvreté. Cela tend à disqualifier l'action des organisations multilatérales de développement dont les mandats sont moins concrets, qui ont pour cible l'ensemble des populations et non les plus pauvres seulement et qui travaillent sur la production de normes et l'amélioration des politiques publiques des États du Sud.

Ces évaluations sont donc un instrument politique, dans la mesure où elles valident certaines approches de l'aide au développement spécifiques à un contexte et une période

[1308] UNEP, *DFID Multilateral Aid Review UNEP Right to Reply*, 2011.
[1309] OIT, *Response by the ILO to the UK-DFID Multilateral Aid Review Report*, 2011.

donnés, et en invalident d'autres, au nom d'une priorité donnée à certains « besoins » des pays en développement ; priorité construite par ces mêmes donateurs qui commanditent ou réalisent les évaluations. Les classements produits entraînent intrinsèquement l'idée qu'il y a des organisations de développement utiles et d'autres inutiles, car ne répondant pas aux priorités actuelles du développement international. Ils empêchent, notamment, de percevoir la complémentarité qui peut naître entre, d'une part les fonds et les banques de développement qui délivrent leur aide dans une logique de verticalité, afin de produire des résultats rapides et massifs et, d'autre part les organisations des Nations unies qui interviennent dans une logique horizontale et travaillent en partenariat avec les gouvernements et les populations du Sud à renforcer leurs capacités institutionnelles, politiques et sociales.

Pour déterminer l'ampleur des effets des évaluations et des classements des donateurs, nous avons analysé le suivi effectué à l'issue de la publication des premiers rapports d'évaluation au milieu des années 2000. Compte tenu du caractère récent de la plupart des processus d'évaluation, seuls le CAD-OCDE, le CGdev (QUODA) et le DFID ont à ce jour établi un tel suivi.

De manière générale, les évaluations montrent que les organismes d'aide font des progrès[1310]. Généralement, ce sont les organisations les moins bien classées qui ont tendance à mener les processus de réformes les plus importants et qui montrent le plus de volonté à améliorer leurs pratiques. Ces organismes semblent avoir une volonté forte de quitter la queue du classement et de rentrer dans la norme des donateurs[1311]. De plus, ce sont les organisations qui ont les plus grandes marges d'évolution.

Ensuite, ce sont les organisations les mieux classées qui montrent les meilleurs progrès. En effet, ce sont les organisations ayant les plus grandes capacités de changement, de par leur fonctionnement efficace et leurs bons résultats originels. Les organisations se situant « dans la norme » sont celles ayant le moins réalisé de progrès sur les indicateurs fixés par les évaluations. À partir de cette analyse, nous émettons l'hypothèse que plus une organisation se situe dans la norme par rapport à ses pairs moins elle est motivée à mettre en place des réformes substantielles. Le fait de se trouver dans la norme inciterait donc à l'immobilisme. Cette hypothèse est généralisable à d'autres organisations et d'autres domaines que l'aide au développement.

[1310] DFID, *Multilateral Review Update, Driving reform to achieve multilateral effectiveness,* décembre 2013 ; BIRDSALL, Nancy, KHARAS, Homi, 2014, *op. cit.*
[1311] *Ibid.*

Les progrès des organisations multilatérales identifiés par les évaluateurs varient significativement selon les secteurs évalués et selon les dispositifs d'évaluation. Cette variabilité montre toute la subjectivité de tels exercices de classement et d'évaluation en l'absence de critères standardisés internationalement, définissant ce qu'est une aide au développement de qualité. Les indicateurs retenus pour les évaluations ne sont pas les mêmes et lorsqu'ils le sont, le seuil à partir duquel une organisation est considérée comme les remplissant de manière satisfaisante diffère. De même, le seuil à partir duquel un progrès est jugé significatif varie selon les critères fixés par les évaluateurs. Ainsi, nous estimons que si ces évaluations constituent un bon indicateur de l'état de fonctionnement et des résultats des organisations multilatérales et participent à l'amélioration de la qualité de l'aide, elles ne fournissent pas une analyse scientifique et objective de la qualité de l'aide.

En termes d'efficience, ou de gestion des ressources financières, notions souvent très liées à la sélectivité de l'aide, QUODA observe peu de progrès[1312]. Les donateurs qui ne prenaient pas en compte la bonne gouvernance comme critère d'allocation de l'aide lors des premières évaluations ne la prennent toujours pas en compte en 2013. Au contraire, le DFID[1313] juge satisfaisants les progrès des donateurs en la matière. Les évaluateurs estiment que toutes les organisations ont mis en place des réformes pour améliorer leurs pratiques d'allocation de l'aide au développement, notamment par la mise en place de procédures de contrôle de la corruption et de la bonne gouvernance des récipiendaires de l'aide. Les Nations unies sont pointées du doigt comme étant celles qui ont fait le moins de progrès en termes d'efficience. Cela est logique, dans la mesure où le mandat des Nations unies a pour principe l'universalité et non la sélectivité.

Promouvoir comme une bonne pratique la sélectivité de l'aide vers les pays les mieux gouvernés et les plus pauvres relève bien d'un choix politique de favoriser les pays dont le mode de gouvernement correspond le mieux aux standards occidentaux. D'autres critères pourraient être choisis, comme le niveau d'inégalité interne dans le pays ou le niveau de fragilité des États. Ces choix ne seraient ni meilleurs ni pires, ils seraient des choix politiques d'allocation de l'aide, tout comme l'est la sélectivité en fonction de la bonne gouvernance. Nous pourrions également imaginer une aide qui n'est pas fondée sur le concept de sélectivité, qu'elle soit politique ou économique, comme prétendent le faire les donateurs émergents. Le fait que les

[1312] QUODA 2014, *op. cit.*
[1313] DFID, 2013, *op. cit.*

évaluations tendent à favoriser une de ces approches de l'aide au développement en fait des instruments profondément politiques.

En ce qui concerne l'aspect de meilleure coordination de l'aide au développement, ou de réduction du fardeau de l'aide pour les pays bénéficiaires, les donateurs réalisent des progrès modérés selon l'évaluation QUODA[1314]. Dans cette évaluation, plus l'aide est concentrée, plus elle est considérée comme efficace. Encore une fois, ce critère est plus politique que technique. Les donateurs choisissent de concentrer leur aide pour des raisons qui relèvent d'abord de l'économie budgétaire dans une période de crise et ensuite des liens politico-économiques qu'ils entretiennent avec les pays du Sud. Ainsi, la France souhaite concentrer son aide majoritairement sur les pays d'Afrique francophone, l'Australie sur les pays de la région du Pacifique-Sud et les États-Unis sur les États qui s'allient à eux dans la lutte contre le terrorisme, ou dans lesquels ils ont engagés des troupes armées. La concentration de l'aide est donc davantage un choix stratégique qu'un choix technique et rationnel. En matière de coordination entre les donateurs sur le terrain, autant l'évaluation QUODA que celle du DFID mettent en avant les bons résultats des organisations multilatérales. En effet, leur nature rend ces organisations plus aptes à se coordonner avec les autres donateurs, car elles ne sont pas soumises à un agenda politique national. En outre, les organisations des Nations unies sont souvent plusieurs à intervenir dans un même pays, ce qui facilite leur tâche de coordination.

Quant à la transparence et au suivi-évaluation, les évaluateurs du QUODA et du DFID ne partagent pas la même analyse des progrès réalisés par les donateurs. Pour les premiers, c'est sur ce critère que les progrès effectués par les donateurs sont les plus grands[1315]. Ils observent que, depuis 2008, un grand nombre de donateurs a adhéré à l'IATI et qu'ils sont « plus méticuleux dans la manière dont ils rapportent les informations sur les activités qu'ils mènent. Les volumes d'aide allant aux pays partenaires en étant assortis de bons cadres d'évaluation et de suivi ont rapidement augmenté »[1316]. Au contraire, le DFID estime que les donateurs ont fait peu de progrès en termes de transparence et de redevabilité[1317].

Concernant la prise en compte du genre par les organisations multilatérales, le DFID évalue les progrès réalisés par les donateurs comme faibles. Le genre ne serait toujours pas une priorité dans les réformes menées. Quant aux donateurs bilatéraux, selon le CAD-OCDE, les progrès réalisés sur cet aspect de la qualité de l'aide seraient plus marqués. Ils relèveraient une

[1314] QUODA 2014, *op. cit.*
[1315] QUODA, 2014, *op. cit.*
[1316] *Ibid.*
[1317] DFID, 2013, *op. cit.*

prise en compte accrue du genre dans les activités d'aide, notamment par les donateurs des pays du Nord de l'Europe et par le Royaume-Uni[1318].

Les évaluateurs lient directement leurs travaux aux réformes mises en œuvre par les donateurs. Nous tenons à nuancer cette corrélation. Les évaluations de la qualité de l'aide se basent en partie, voire totalement, sur des évaluations internes déjà réalisées par les organisations. En conséquence, les forces et faiblesses identifiées par ces évaluations externes ont souvent déjà été relevées par des évaluations internes et les réformes afférentes engagées. La valeur ajoutée principale des évaluations externes tient surtout dans le bilan global qu'elles réalisent sur la qualité et, dans une moindre mesure, l'efficacité de l'aide. Elles contribuent également à rendre plus transparente l'aide au développement, une politique publique jusque-là très opaque. Par ces mécanismes internationaux, les organismes d'aide sont incités à publier des informations de plus en plus précises et exhaustives sur leurs activités. Les évaluations produisent également des classements, explicites ou implicites, qui font émerger une forme de concurrence ou d'émulation entre les acteurs de l'aide.

- **Conclusion de chapitre**

Nous avons analysé, dans ce chapitre un phénomène récent et peu étudié consistant dans la multiplication des procédures d'évaluation, par les pairs et par des centres de recherche, de la qualité et de l'efficacité de l'aide des donateurs. Selon nous, ce phénomène traduit un besoin grandissant des donateurs de prouver que leur mandat est légitime et qu'ils contribuent efficacement à améliorer l'état de développement des pays du Sud. Il est aussi le résultat d'un plaidoyer fort de la société civile internationale pour que les politiques d'aide des donateurs bilatéraux et multilatéraux soient soumises à davantage de transparence.

Ce chapitre a, encore une fois, démontré que les évaluations sont un outil politique. En effet, lorsque nous analysons les critères d'évaluation créés pour juger la qualité de l'aide au développement, nous comprenons qu'ils contribuent à légitimer et à perpétuer les approches dominantes actuelles de l'aide au développement. Ces évaluations ont donc une fonction de diffusion des bonnes pratiques et de légitimation du régime actuel de l'aide au développement.

Le caractère politique de ces évaluations se voit également dans l'évolution des critères les composant, qui reflètent les changements idéologiques de l'aide. Ainsi, depuis que le

[1318] OECD, 2009, *op. cit.*

Partenariat mondial pour le développement a succédé à l'Agenda sur l'efficacité de l'aide, les évaluateurs modifient leurs méthodologies d'évaluation, afin de mieux être en phase avec cette nouvelle doctrine. Ce constat n'est pas une critique négative, il permet seulement de rompre avec l'idée selon laquelle les évaluations sont des outils désincarnés et coupés du contexte politique dans lequel elles sont mobilisées.

Enfin, nous avons démontré que ces évaluations reclassent les donateurs selon leurs performances. Elles contribuent ainsi à instaurer un phénomène de concurrence dans le système de l'aide, le faisant de plus en plus ressembler à un marché, où les bailleurs de fonds sont des financeurs qui aiguillent leurs « investissements » vers les « clients » les plus « rentables ». Aujourd'hui, l'enjeu pour les donateurs, et en premier lieu pour les organisations multilatérales, est de prouver qu'elles sont performantes ou qu'elles apportent de la *value for money*. Les organisations déclassées par les évaluations et perçues comme non rentables par les donateurs bilatéraux sont, en conséquence, délaissées et survivent avec des fonds dérisoires, situation qui renforce leur difficulté à se réformer et à produire des résultats.

Toutefois, cette inefficacité que reflètent les évaluations n'est pas objective. Elle est le symptôme d'un système de l'aide qui favorise un certain type d'efficacité, l'atteinte des OMD, et qui porte de moins en moins d'intérêt à la production de normes, aux actions de plaidoyer et à la formation (des personnels de santé par exemple). Par ailleurs, les reclassements que nous avons analysés tendent à entraîner une harmonisation des pratiques de l'aide, les organisations de développement déclassées cherchant à se rapprocher de la norme, même si cette norme a peu de sens par rapport à leurs mandats.

Conclusion

Notre recherche est née de questionnements quant à la construction de « bonnes pratiques internationales » d'évaluation des programmes d'aide au développement. Retracer l'histoire de ces bonnes pratiques nous a permis de mettre en avant la nature politique de cet objet d'étude. En effet, cette analyse nous a permis d'observer que les approches et pratiques de l'évaluation ont évolué en fonction des changements dans les politiques d'aide au développement. Aborder l'évaluation de l'aide au développement d'un point de vue « macro », nous a permis de démontrer que tous les projets et programmes d'aide sont évalués selon un ensemble de critères standardisés, quel que soit le secteur ou le pays dans lequel ils sont mis en œuvre. Nous avons ensuite pu étudier les bénéfices de l'institutionnalisation d'un tel cadre normatif, ainsi que ses limites et les difficultés qu'elle peut entraîner.

Nous avons observé que les changements dans les pratiques et les usages de l'évaluation suivent un cycle. Dans les années 1960 et 1970, l'évaluation était utilisée à des fins d'apprentissage pour tirer des leçons des projets expérimentaux. Dans les années 1980, après le tournant néolibéral des politiques d'aide au développement, les approches quantitatives de l'évaluation furent privilégiées pour mesurer l'efficience et l'efficacité des projets d'aide, ainsi que pour contrôler que les pays en développement menaient bien les réformes « recommandées » par le Consensus de Washington.

Au milieu des années 1990, après le constat de l'échec des politiques d'aide des années 1980, qualifiées de « décennie perdue de l'aide au développement », les organisations multilatérales développèrent une approche sociale de l'aide au développement qui mettait l'accent sur la lutte contre la pauvreté, le renforcement des systèmes éducatifs et de santé et la promotion de la bonne gouvernance dans les États du Sud. Alors, les évaluateurs mirent en place des approches qualitatives et participatives de l'évaluation des programmes d'aide pour mieux renseigner les progrès des bénéficiaires sur ces aspects immatériels du développement.

Toutefois, parallèlement, les politiques publiques des États donateurs étaient influencées par le courant du New Public Management, qui donnait à l'évaluation une fonction de redevabilité permettant de contrôler l'efficacité organisationnelle et les résultats des administrations publiques. La production de résultats chiffrés réguliers devint un gage de bonne gouvernance des politiques publiques. À mesure que le New Public Management se diffusa

dans les organismes d'aide au développement, l'approche gestionnaire de l'évaluation servant un objectif de redevabilité prit le pas sur les approches qualitatives et participatives.

Cette approche gestionnaire de l'évaluation se renforça dans les années 2000, pour finalement être remise en cause à la fin de la décennie en raison d'une pression des bénéficiaires de l'aide et des ONG internationales pour que le système de la coopération pour le développement soit refondé vers davantage de partenariat et de redevabilité mutuelle entre donateurs traditionnels, donateurs émergents, bénéficiaires de l'aide, ONG et secteur privé. Dès lors, les donateurs mirent de nouveau en avant une approche participative de l'évaluation pouvant contribuer à rendre plus égales et collaboratives les relations entre ces différents acteurs de l'aide, permettant aux bénéficiaires de se réapproprier les programmes d'aide et de rendre les donateurs plus redevables de la qualité de l'aide qu'ils leur fournissent.

Ainsi, nous avons démontré que la pratique de l'évaluation oscille en permanence entre une fonction de redevabilité et une fonction d'apprentissage, ainsi qu'entre des méthodologies dites scientifiques et des méthodologies participatives. Elles sont alternativement favorisées par les évaluateurs selon les besoins des donateurs et les contraintes organisationnelles et contextuelles dans lesquelles ils interviennent. Nous montrons donc que l'évaluation n'est pas un instrument indépendant et qu'elle est déterminée par les approches politiques de l'aide. Toutefois, étant donné que l'évaluation est devenue un pilier incontournable de l'aide, les donateurs sont pris dans une nécessité croissante de redevabilité et d'évaluation et en font un élément central de leurs politiques et de leurs pratiques. Avoir des évaluations aux résultats positifs est devenu un souci primordial pour les organismes d'aide au développement, qui ont alors tendance à privilégier les activités et les publics qui apporteront des résultats favorables.

Nous avons alors cherché à comprendre pourquoi les fonctions d'apprentissage et de redevabilité étaient autant considérées comme antinomiques par les acteurs de l'aide. Notre recherche a démontré que les décideurs, managers et opérationnels des organismes d'aide estiment que les évaluateurs ne peuvent pas à la fois contrôler les pratiques de leurs collègues à des fins de redevabilité et être les facilitateurs d'un apprentissage interne. Se sentant jugés et poussés à la performance, le personnel des agences d'aide estime ne pas être dans de bonnes conditions pour apprendre des évaluations. La fonction d'apprentissage tend donc à s'effacer alors qu'elle est une des caractéristiques qui distingue l'évaluation des disciplines voisines de l'audit et du contrôle interne.

De plus, nous avons montré que l'expertise des évaluateurs est peu reconnue dans le secteur de l'aide au développement, même s'ils ont, de par leur positionnement dans les

organisations, une légitimité « hiérarchique ». Au demeurant, les donateurs, même s'ils ont de plus en plus recours à l'évaluation, restent peu sensibilisés à leur utilité et peu incités à s'informer sur les conclusions des rapports d'évaluation. Ainsi, les évaluateurs s'engagent actuellement dans des initiatives de plaidoyer dans les organismes d'aide pour convaincre les dirigeants et les opérationnels de l'utilité des évaluations au-delà de la mesure des résultats.

Notre recherche a permis de comprendre que ce manque de légitimité des évaluateurs dans les organismes d'aide est dû au fait que les dirigeants des agences d'aide ont cru qu'il suffirait de mettre en place des réformes organisationnelles, en plaçant les unités d'évaluation à un échelon hiérarchique élevé des agences d'aide, pour qu'une culture de l'évaluation se diffuse dans leurs organisations. Or, à travers les exemples de l'AFD et de la Banque mondiale, nous avons montré que c'est la présence d'acteurs investis dans la promotion de l'évaluation qui permet un essor de la légitimité et de l'influence de l'évaluation dans les organismes d'aide. Dès qu'il n'y a plus d'acteurs forts pour soutenir le développement de la fonction d'évaluation, cette dernière tend à perdre en influence. Si les postes de directeurs des unités d'évaluation sont relégués à des individus peu formés à l'évaluation et peu convaincus de son utilité, ce qui est souvent le cas, alors la culture organisationnelle de l'évaluation reste limitée. Un autre facteur fondamental de la diffusion d'une culture de l'évaluation est l'implication du personnel des agences d'aide dans la mise en œuvre des évaluations. Toutefois, il manque des incitations pour encourager cette participation qui est chronophage.

Bien que la culture de l'évaluation et de l'apprentissage reste peu développée dans un grand nombre d'organisme d'aide, nous avons remis en cause l'idée selon laquelle les évaluations ne seraient pas utilisées. Nous avons montré que si les évaluations n'avaient pas un usage instrumental menant, directement après leur publication, à des réformes de grande envergure, elles ont un usage diffus qui se concrétise sur le long terme. L'usage direct des évaluations se produit surtout au niveau opérationnel lorsqu'elles apportent des solutions concrètes à des problèmes se posant aux responsables de projets. Au niveau international, l'usage direct se fait surtout sous la forme de plaidoyers des décideurs se basant sur les conclusions des évaluations pour défendre un projet ou un programme qui leur tient à cœur.

La perception, répandue dans les agences d'aide, selon laquelle l'évaluation est peu utile provient d'attentes inappropriées en ce qui concerne les effets de l'évaluation. L'évaluation ne peut pas apporter de changements massifs et brutaux au niveau des stratégies et des politiques d'aide. En effet, contrairement à ce qui est présenté dans la littérature grise et dans les discours des dirigeants des organismes d'aide au développement, l'évaluation n'est pas un élément

central de la prise de décision. Elle constitue un élément parmi d'autres, qui sont parfois plus puissants, notamment lorsqu'ils sont de l'ordre de la défense des intérêts diplomatiques des pays donateurs. L'évaluation n'a donc pas dépolitisé l'aide au développement.

Toujours sur le thème de la légitimation de l'expertise de l'évaluation, nous avons démontré qu'elle passe, pour les évaluateurs du Nord, par la création de compétences en évaluation au Sud. En effet, dans un système de l'aide où, depuis le milieu des années 1990, l'accent est mis sur le fait que les projets et les programmes doivent être mis en œuvre de manière partenariale avec les pays bénéficiaires, il faut que ces derniers soient représentés dans les processus d'évaluation. Ainsi, pour que leur expertise apparaisse légitime et crédible, les évaluateurs du Nord doivent s'allier avec des experts du Sud pour réaliser les évaluations. Or, pour que les experts et gouvernements du Sud puissent participer activement aux évaluations, il faut qu'ils soient formés à cette activité.

Notre recherche a démontré que les formations réalisées par les donateurs pour transférer des compétences en évaluation au Sud ont été, dans la plupart des cas, des échecs. Alors, ces formations académiques ont été substituées par des activités de *learning by doing*, au cours desquelles les acteurs du Sud se forment à l'évaluation en la pratiquant avec les donateurs. Nous avons montré, qu'en théorie, cette approche de la diffusion d'une expertise de l'évaluation au Sud semblait meilleure. Toutefois, l'évaluation des programmes d'aide est toujours conçue par les donateurs pour répondre à leurs besoins de redevabilité nationaux. Ainsi, les experts du Sud engagés pour mener les évaluations aux côtés d'évaluateurs du Nord sont peu intéressés par ces exercices et par le gain de compétences qu'ils pourraient en retirer. Les évaluateurs du Nord les relèguent alors, par convenance ou par dépit, à des positions de subalternes, les utilisant comme des facilitateurs pour organiser des entretiens ou mieux comprendre le contexte politique et culturel des pays bénéficiaires.

De plus, nous avons prouvé que former quelques experts en évaluation au sein des gouvernements des pays du Sud ne suffit pas pour diffuser une culture d'évaluation dans ces pays. Si ces derniers ont mis en place des structures d'évaluation pour répondre aux exigences des donateurs en termes de bonne gouvernance, elles restent aujourd'hui des coquilles vides en raison d'un manque d'intérêt de la classe politique pour l'évaluation ou d'un manque de moyens humains et financiers. Nous avons démontré que les résistances locales au développement de la fonction d'évaluation sont dues au fait que cette dernière est vue par les gouvernements comme étant un instrument porteur de changements politiques profonds impliquant une amélioration de la performance et de la transparence des politiques publiques, une amélioration

de la démocratie participative, ou encore une baisse de la corruption. Dans des contextes politiques qui ne sont pas favorables à ces changements, l'introduction de systèmes d'évaluation reste sans résonnance. Une autre source de résistance est que, dans les pays du Sud, l'évaluation est vue comme un objet participant à une domination symbolique des donateurs, qui sont perçus comme imposant par ce biais leurs principes de bonne gouvernance.

Les évaluateurs des organismes d'aide sont conscients du fait que l'évaluation est marquée par l'image d'une domination du Nord par rapport au Sud dans le cadre de l'aide au développement. Quelques évaluateurs des Nations unies, notamment de l'UNICEF, ont alors eu l'idée de mettre en place un mécanisme non plus vertical de transfert des compétences en évaluation, mais horizontal, ou « Sud-Sud », en créant des réseaux d'évaluateurs nationaux et régionaux au Sud. Ils ont alors sélectionné des « champions » de l'évaluation dans les pays du Sud et les ont aidés à créer leurs associations d'évaluateurs, afin de diffuser une culture d'évaluation dans leurs pays. À ces associations nationales s'ajoutent des réseaux régionaux qui permettent un apprentissage « Sud-Sud » de l'évaluation. Enfin, les sociétés internationales d'évaluateurs mises en place donnent une visibilité internationale aux évaluateurs du Sud et leur permettent de se créer un réseau professionnels. Notre recherche a permis de découvrir qu'au-delà de l'enjeu de la formation d'évaluateurs au Sud, la création de réseaux relève d'une volonté des évaluateurs de se professionnaliser et de se constituer en une forme de communauté épistémique, afin d'augmenter leur influence politique sur les organismes de développement et les États du Sud.

Toutefois, nous avons montré que les réseaux d'évaluateurs du Sud se retrouvent souvent démunis une fois que les organisations multilatérales se retirent pour les laisser se développer de manière autonome. Faute d'un nombre suffisant de membres et de moyens financiers et d'influence, ces réseaux connaissent des difficultés à rentrer en relation avec les instances gouvernementales de leurs pays. En conséquence, leur pouvoir de changement sur les politiques publiques est faible et ils peinent à diffuser une culture de l'évaluation dans leurs pays. Ils ont besoin de passer par les réseaux régionaux et internationaux pour gagner cette influence. De plus, nous avons constaté que la création de réseaux internationaux n'a pas permis aujourd'hui d'instaurer des relations professionnelles pérennes entre la plupart des évaluateurs du Nord et du Sud. Les réseaux de ces deux espaces du système international de l'aide interagissent peu, et les rencontres permises par les séminaires annuels des réseaux internationaux ne dépassent souvent pas le stade de l'échange de contacts. Toutefois, ces réseaux étant, pour la plupart, très récents, il est encore trop tôt pour juger de manière définitive de leur capacité à créer une

communauté internationale d'évaluateurs et à devenir des « opérateurs de changement » au niveau national ou international.

Dans cette deuxième partie, nous avons montré que l'évaluation est au cœur de relations de pouvoir entre les acteurs de l'aide. Donner une plus grande légitimité aux évaluateurs, c'est augmenter leur capacité de changement sur les politiques d'aide au développement. Créer des compétences d'évaluation au Sud revient, pour les donateurs, à perdre le contrôle sur le jugement de l'efficacité de leurs programmes d'aide. Pour les gouvernements du Sud, c'est accepter que les fragilités de leurs régimes soient révélées, ce qui constitue un effet déstabilisateur potentiel. Ainsi, des résistances fortes apparaissent face à la reconnaissance de la légitimité des évaluateurs en tant qu'acteurs contribuant légitimement au changement et à l'apprentissage. Les cantonner à un statut de techniciens, dont la compétence se limite à mesurer les résultats des programmes d'aide, garantit plus de stabilité politique. Nous avons, enfin, constaté que ces résistantes proviennent d'une inquiétude, aussi bien des bénéficiaires de l'aide que du personnel des agences d'aide, quant aux remises en cause des pratiques existantes ou aux effets de déclassement que les évaluations les concernant pourraient entraîner. En effet, évaluer suppose l'établissement d'une échelle de valeurs positives et négatives, à laquelle l'objet évalué se trouve ensuite confronté.

Dans une troisième partie, nous avons analysé ces effets de déclassement. Nous avons découvert qu'ils étaient variables selon les espaces du système de l'aide au développement. Alors que les donateurs émergents, qui affichent un rejet des « bonnes pratiques » d'efficacité et d'évaluation, sont déclassés sur la scène internationale et pointés du doigt par les donateurs traditionnels comme étant des mauvais donateurs, ils se retrouvent reclassés dans les pays en développement, où les gouvernements apprécient leur aide moins bureaucratique.

Ne pouvant plus ignorer les donateurs émergents et leurs conceptions alternatives de l'aide, les donateurs traditionnels ont dû faire évoluer leur conception de l'efficacité de l'aide pour se rapprocher de celle des donateurs émergents. Ainsi, ils ont mis en avant des bonnes pratiques prônant des relations plus égalitaires avec les bénéficiaires de l'aide, telles que la redevabilité mutuelle, la mise en œuvre partenariale de l'aide, les évaluations participatives, etc. De leur côté, les donateurs émergents, se positionnant en dehors du régime normatif actuel de l'aide au développement, sont influencés par ses bonnes pratiques et commencent à développer des fonctions évaluatives et à adopter des procédures de transparence. L'influence entre donateurs émergents et traditionnels est donc réciproque.

De plus, nous avons vu que les oppositions politiques qui se créent autour de l'efficacité de l'aide et de l'évaluation sont plus complexes qu'une simple dichotomie entre donateurs émergents et donateurs traditionnels. Des désaccords apparaissent également entre ces derniers. En effet, des donateurs traditionnels déclarent se sentir proches de la « culture de l'aide » des donateurs émergents. C'est ainsi, par exemple, que le Japon a développé le concept d'une « culture asiatique de l'aide au développement » qui regrouperait sous des caractéristiques communes tous les donateurs asiatiques traditionnels et émergents.

Un autre effet de déclassement que nous avons exploré est celui d'une possible éviction de l'accès à l'aide au développement des pays du Sud les moins performants. Nous avons trouvé que, contrairement à notre hypothèse de départ, les évaluations portant sur l'efficacité des programmes d'aide ont peu d'effet sur l'allocation de l'aide. Ainsi, ce n'est pas parce que, dans un pays, les programmes mis en place ont des résultats jugés médiocres que ce pays verra son accès à l'aide limité. Nous avons démontré que c'est plutôt l'évaluation de la gouvernance des pays du Sud qui influence les montants d'aide qu'ils reçoivent. Ce sont les pays jugés comme les moins bien gouvernés qui se retrouvent déclassés, même lorsqu'ils appartiennent à la catégorie des pays les moins avancés.

À partir de ce constat, nous avons démontré que l'évaluation sert de plus en plus des fins de gouvernementalité des pays et des populations jugés à risque. Les donateurs se servent du suivi-évaluation pour rassembler des données sur les populations les plus vulnérables et normaliser leur comportement selon les indicateurs et objectifs fixés. L'évaluation devient ainsi un instrument de contrôle de l'ordre social dans les États à faible gouvernance et, plus globalement, sert à la logique actuelle de sécurisation de l'aide au développement. Par ailleurs, plus les donateurs sont soumis à une pression de produire des résultats de court terme, plus ils sont contraints à imposer des réformes aux pays du Sud et donc à exercer une domination symbolique. Dans les pays les moins avancés, les gouvernements n'ont souvent pas la volonté ou la capacité d'interférer avec l'entreprise de gouvernementalité des donateurs et se satisfont de cette situation, de laquelle ils tirent divers bénéfices : maintien des fonds d'aide, possibilité de se tourner vers d'autres problématiques de gouvernance, évitement de la mise en place de réformes déstabilisatrices pour leur maintien au pouvoir, etc. De plus, il existe des cas où la publication d'évaluation a permis d'attirer l'attention de la « communauté des donateurs » sur des difficultés éprouvées par les populations du Sud, qui restaient jusqu'alors inconnues ou sous-estimées. Ainsi, interpréter ce phénomène comme une domination volontaire des donateurs sur les bénéficiaires est réducteur.

Seuls les gouvernements des pays à revenus intermédiaires parviennent à se positionner comme de réels partenaires des donateurs dans l'évaluation des programmes d'aide au développement. Selon nous, le potentiel émancipateur et égalitaire de l'évaluation ne peut se réaliser que dans des pays où les structures gouvernementales sont solides, où la société civile est organisée et capable de négocier avec les donateurs, et où les dirigeants ont assez de marge de manœuvre diplomatique pour faire entendre leurs voix aux donateurs dans la formulation des indicateurs et objectifs qui déterminent les exercices d'évaluation.

Nous avons observé que pour redonner du pouvoir aux gouvernements bénéficiaires de l'aide, le CAD-OCDE a instauré la « bonne pratique » de la redevabilité mutuelle, selon laquelle les bénéficiaires de l'aide seraient légitimes pour évaluer la qualité de l'aide des donateurs. Toutefois, nous avons démontré qu'au vu des faibles capacités en évaluation des autorités publiques des pays du Sud et de leur dépendance aux donateurs, les rares mécanismes d'évaluation mis en place pour leur permettre de tenir redevables les donateurs ont peu d'effets transformateurs sur les pratiques de ces derniers. Nous avons souligné que le fait d'importer de tels mécanismes techniques « d'égalitarisation » des relations entre les donateurs et les bénéficiaires ne peut être inefficace que s'ils sont accompagnés de changements politiques dans le régime de l'aide, renouvelant les rapports de force entre donateurs et bénéficiaires. En effet, le fait que les bénéficiaires de l'aide soient légitimes d'exprimer des critiques sur l'aide fournie par les donateurs est encore une idée controversée, « derrière la scène », par certains donateurs, notamment les États-Unis.

Nous avons également introduit l'idée selon laquelle le fait que les donateurs ont longtemps été peu soumis à une exigence de redevabilité envers les autres acteurs de l'aide (bénéficiaires ou autres donateurs) a renforcé l'image de l'évaluation comme étant un outil de contrôle des donateurs sur les pays du Sud. Cette absence de redevabilité internationale pesant sur les donateurs n'est plus acceptable dans un système de l'aide qui se veut partenarial et fondé sur le principe de redevabilité mutuelle. Pour pallier ce manque, des mécanismes d'évaluation de la qualité des politiques d'aide des donateurs, par d'autres donateurs ou par des centres de recherche, ont été créés ces dernières années. Toutefois, nous avons montré que la plupart de ces évaluations se concentrent sur les organisations multilatérales. En effet, dans un contexte où les donateurs doivent produire des preuves de résultats sur toutes leurs activités, il était problématique, pour eux, de ne pas avoir de moyens de suivre et d'évaluer l'utilisation des fonds qu'ils versent aux organisations multilatérales.

Certaines de ces évaluations produisent des classements des donateurs, afin d'inciter les « mauvais élèves » à améliorer leur efficacité organisationnelle et leurs résultats sur le terrain. Nous avons montré les organismes des Nations unies sont systématiquement déclassés, ce qui tend à prouver que les évaluations sont biaisées en faveur d'une conception politique de l'aide au développement qui favorise les résultats quantifiables et de court terme, au détriment d'actions portant sur la production de normes, le renforcement des institutions, le conseil politique, etc. Les classements favorisent également les organisations qui se concentrent sur les OMD, telles que le FMSTP, au détriment des organisations dont le mandat s'éloigne de ces objectifs, par exemple l'OIT ou l'UNESCO.

L'analyse de ces évaluations de l'efficacité des politiques d'aide des donateurs nous a permis d'observer qu'elles ne parviennent pas à donner une appréciation au niveau national ou international de l'efficacité de l'aide. Elles se cantonnent souvent à juger de l'efficacité organisationnelles des organismes d'aide, ou de leurs résultats sur un échantillon restreint de programmes. Ainsi, les débats sur l'efficacité de l'aide divisent-ils autant les experts de l'aide au développement aujourd'hui qu'il y a trente ans.

Ainsi, notre recherche a montré que les usages et les effets de l'évaluation sur les politiques et les pratiques d'aide au développement sont multiples et complexes. Nous en proposons une typologie :

- les évaluations à des fins de redevabilité, qui sont utilisées par les donateurs pour prouver leurs résultats et l'efficacité de leur action dans les pays du Sud ;
- les évaluations à des fins d'apprentissage, qui servent aux donateurs et aux gouvernements bénéficiaires à améliorer les projets et programmes qu'ils mettent en œuvre pour qu'ils gagnent en efficacité ;
- les évaluations « émancipatrices », qui ont pour fonction de donner la parole aux bénéficiaires, de rendre plus égalitaires leurs relations avec les donateurs et de leur permettre à de gagner des compétences en évaluations ;
- les évaluations « de plaidoyer », pour soutenir les décisions politiques des responsables des agences d'aide ou de leurs tutelles financières ;
- les évaluations à des fins de contrôle, mises en place par les donateurs pour garder la main sur les programmes de développement qui sont mis en place dans les pays du Sud, notamment ceux qui prennent la forme « d'aide budgétaire »[1319] ;

[1319] Pour une analyse critique de la mise en œuvre du principe d'appropriation par les donateurs, voir : WITHFIELD, Lindsay, *The Politics of Aid : African Strategies for Dealing with Donors*, Oxford University Press, 2009, 401p.

- l'évaluation à des fins de « gouvernementalité », pour maîtriser et influencer les trajectoires de vie des populations vulnérables des pays bénéficiaires, afin de limiter les risques d'échec des programmes d'aide et de sécuriser le système international ;
- les évaluations à des fins mobilisatrices, qui sont utilisées pour susciter une prise de conscience internationale sur de nouveaux enjeux en leur donnant une visibilité.

Une évaluation peut recouvrir plusieurs de ces fonctions qui ne s'excluent pas l'une l'autre. Toutefois, certaines d'entre-elles apparaissent contradictoires, par exemple l'apprentissage et la redevabilité, ou le contrôle et l'émancipation. L'indétermination entre ces différents usages fait de l'évaluation un instrument de gouvernement ambivalent. Cette ambivalence reflète, selon nous, une politique internationale de l'aide qui ne parvient pas à passer d'une approche tutélaire à une approche émancipatrice envers les bénéficiaires de l'aide. Si la seconde est privilégiée dans les discours, les pratiques restent ancrées dans la première en raison d'intérêts politiques et de contraintes contextuelles diverses.

Malgré les critiques qui peuvent être formulées quant à la manière dont est actuellement mise en œuvre l'évaluation par les donateurs, il advient de reconnaître qu'elle constitue un progrès important de l'aide au développement. En effet, l'évaluation permet une meilleure transparence des politiques d'aide et l'accumulation de données quantitatives et d'informations qualitatives sur les actions de terrain. De plus, l'évaluation entraîne un plus grand souci des donateurs quant aux résultats de leurs programmes d'aide.

En ce qui concerne les populations bénéficiaires, l'évaluation permet souvent de donner de la visibilité à leurs problématiques de développement, d'autant plus maintenant qu'elles sont presque systématiquement publiées. Si l'on peut critiquer la surenchère évaluative que connaît aujourd'hui l'aide au développement, il faut reconnaître avant tout que l'évaluation favorise la mise en œuvre de programmes d'aide de qualité et éthiques. Ainsi, un retour en arrière apparaît peu souhaitable. L'enjeu est plutôt que les donateurs et les bénéficiaires parviennent à prendre davantage de recul par rapport aux travaux d'évaluation, afin de leur donner du sens et d'en tirer des leçons sur les expériences évaluées. Le défi est également de redonner à l'évaluation la capacité de réaliser son potentiel d'amélioration du dialogue et de la compréhension et de l'apprentissage mutuels entre les différents acteurs de l'aide.

Notre recherche se concentrant sur l'évaluation dans le cadre des organismes gouvernementaux d'aide au développement, une autre piste de réflexion pourrait être l'étude

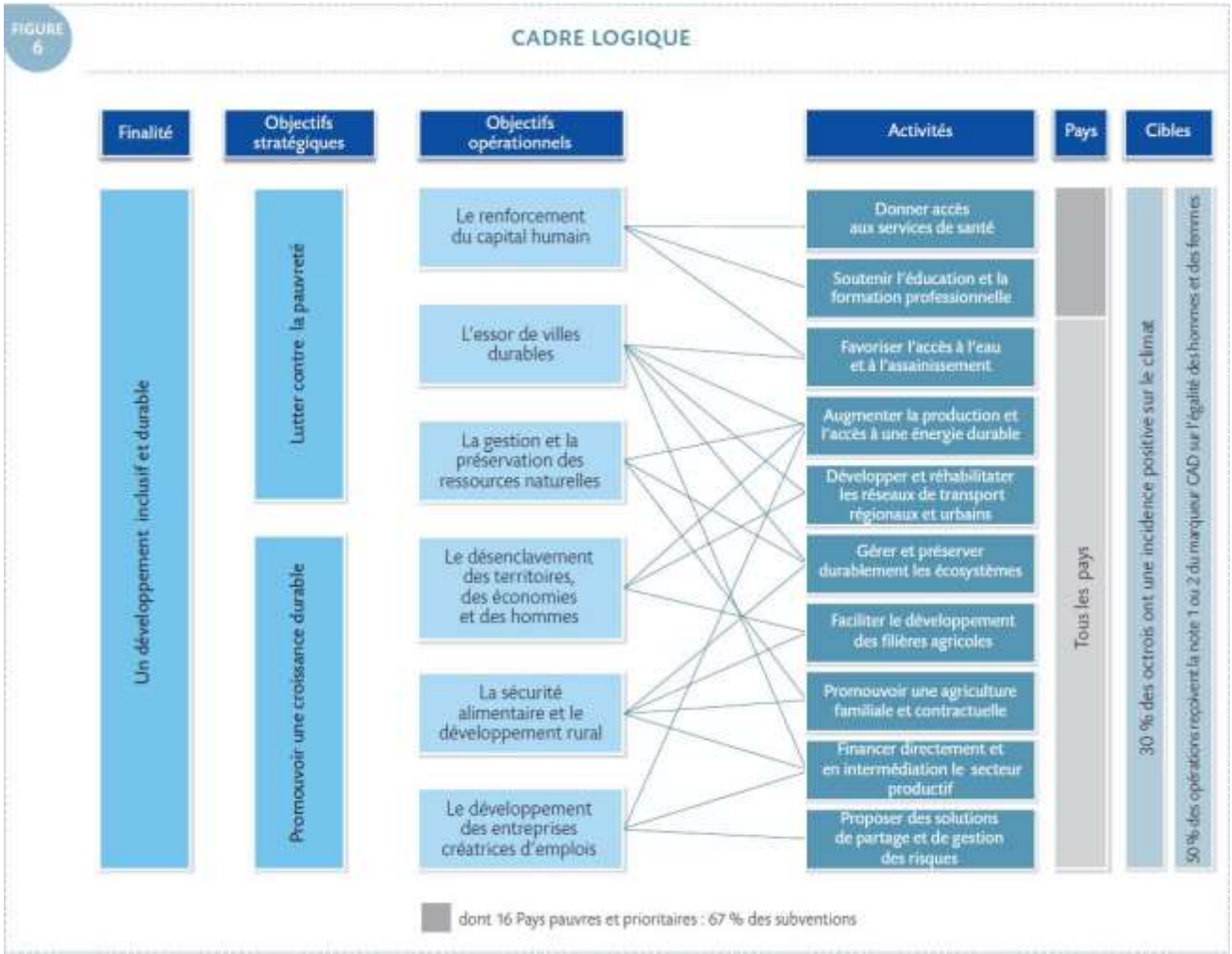
des pratiques et des usages de l'évaluation développés par les ONG dans le cadre de l'aide humanitaire. Ces organisations connaissent-elles les mêmes contraintes de redevabilité que les donateurs gouvernementaux et, dans ce cas, utilisent-elles l'évaluation pour répondre à cette obligation plutôt qu'à des fins d'apprentissage ? Les contextes d'action spécifiques des ONG modifient-ils leurs conceptions de l'évaluation et la sociologie des acteurs impliqués dans les évaluations ?

Une autre piste de recherche pourrait être d'analyser les causes de la montée en puissance de l'évaluation *ex ante* dans le secteur de l'aide au développement et le poids qu'ont ces évaluations sur la planification des programmes d'aide. En effet, nous avons perçu dans cette recherche que la hausse des procédures d'évaluations ex ante, notamment des risques environnementaux et sociaux des projets et programmes, par les donateurs entraînaient des questionnements stratégiques importants sur le terrain de l'aide. Les donateurs semblent avoir du mal à trouver un équilibre entre, d'une part la mise en place d'évaluations *ex ante* et de mesures de contrôle des risques exhaustives et, d'autre par le maintien de l'efficience des projets d'aide, sans parler des tensions que ces évaluations peuvent entraîner avec les gouvernements bénéficiaires.

Enfin, l'analyse de l'évaluation que nous avons appliquée à l'aide au développement pourrait être utilisée dans le cadre d'autres secteurs et d'auteurs acteurs des politiques publiques : la santé ou l'éducation par exemple. Il s'agirait d'analyser la sociogenèse de l'évaluation dans ces secteurs de politiques publiques et d'explorer les transformations qu'elle entraîne sur les pratiques, les résultats recherchés, les relations entre acteurs, la performance des projets et programmes, etc. Cela pourrait également permettre de soulever la question de savoir pourquoi les Cours des comptes européennes se saisissent peu du mandat d'évaluation indépendante des politiques publiques, laissant cette tâche à des cabinets de consultants privés.

Annexe 1 : Cadre logique

Exemple de cadre logique de l'AFD représentant la stratégie d'intervention générale de l'agence en Afrique Sub-saharienne.



Source : site Internet de l'AFD : http://www.afd.fr/home/pays/afrique/strategie-afr, consulté le 27 décembre 2014

Annexe 2 : Approches de l'aide et conditionnalités sous-jacentes

Rôle de l'aide	Origine de la pauvreté	Allocation de l'aide	Principe d'action	Principe de justice	Temporalité
Aide comme un droit	Diagnostic "situationnel" : les pays ne portent pas la responsabilité de leur pauvreté.	Allocation « automatique » en fonction du besoin des bénéficiaires et de critères universels d'éligibilité.	Principe libéral : pas de limitation des choix et libertés des bénéficiaires.	Justice "distributive" ou "procédurale : démarche déontologique qui se justifie par le besoin de corriger une répartition des ressources qui n'est pas légitime du point de vue des droits et/ou des mérites et/ou des besoins de chacun.	Aide à vocation permanente.
Aide incitative	Diagnostic « dispositionnel » : les pays portent la responsabilité de leur pauvreté.	Allocation "conditionnelle" sur la base d'un contrat avec le bénéficiaire qui s'engage à conduire un certain nombre d'actions.	Principe perfectionniste ou développemental : le donateur promeut un modèle de développement dans un souci de rattrapage.	Justice " corrective " : démarche conséquentialiste qui répond au souci de la réparation d'injustices situées dans le temps. Elle repose sur l'évaluation du préjudice puis sur sa réparation.	Aide temporaire.
Aide comme soutien sélectif	Diagnostic « dispositionnel » : les pays portent la responsabilité de leur pauvreté.	Allocation "sélective" à un certain nombre de bénéficiaires afin qu'ils réalisent des actions et des objectifs selon des modalités qu'ils ont eux-mêmes choisies. L'allocation est fondée sur l'évaluation des besoins et de la capacité des bénéficiaires à mettre en places lesdites actions.	Principe libéral par nature mais les critères d'évaluation peuvent se révéler perfectionnistes s'ils influencent les politiques des bénéficiaires.	Justice comme" équité": démarche liée à la perception d'un niveau minimal de droits, de considération, de ressources ou de bien-être dus à tous les êtres humains. La finalité de l'aide est, selon les critères retenus, de permettre à tous de dépasser un niveau de privation considéré comme inacceptable.	Système temporaire si cessation de l'aide lorsque les finalités sont atteinte, permanent si elles sont revues à la hausse une fois réalisées.

Sources : Réalisé à partir de COGNEAU, David, NAUDET, Jean-David, *Qui mérite l'aide ? Égalité des chances versus sélectivité*, AFD, 2004

Annexe 3 : Revue des principales études académiques sur l'efficacité de l'aide (1970-2012)

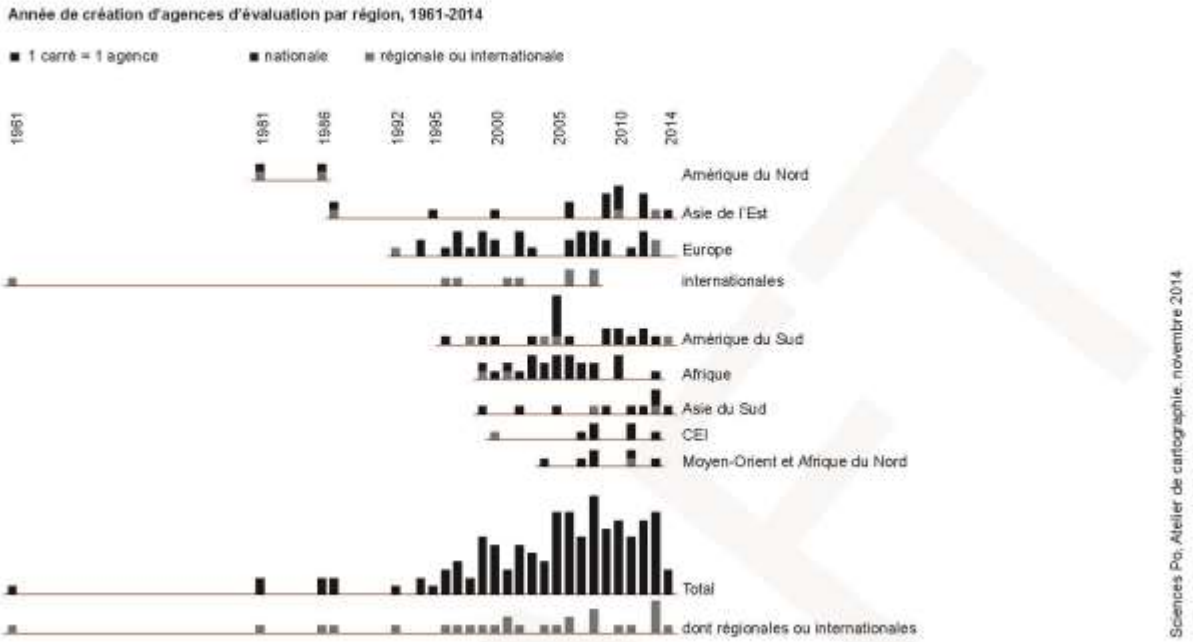
Ce tableau permet d'avoir une vision d'ensemble sur les évolutions de la perception de la notion « d'efficacité » de l'aide par les chercheurs sur l'aide au développement. Il permet de comprendre, à travers les décennies, quels étaient les critères d'évaluation des chercheurs pour statuer de l'efficacité ou de l'impuissance de l'aide au développement pour contribuer à la croissance économique des pays bénéficiaires de l'aide :

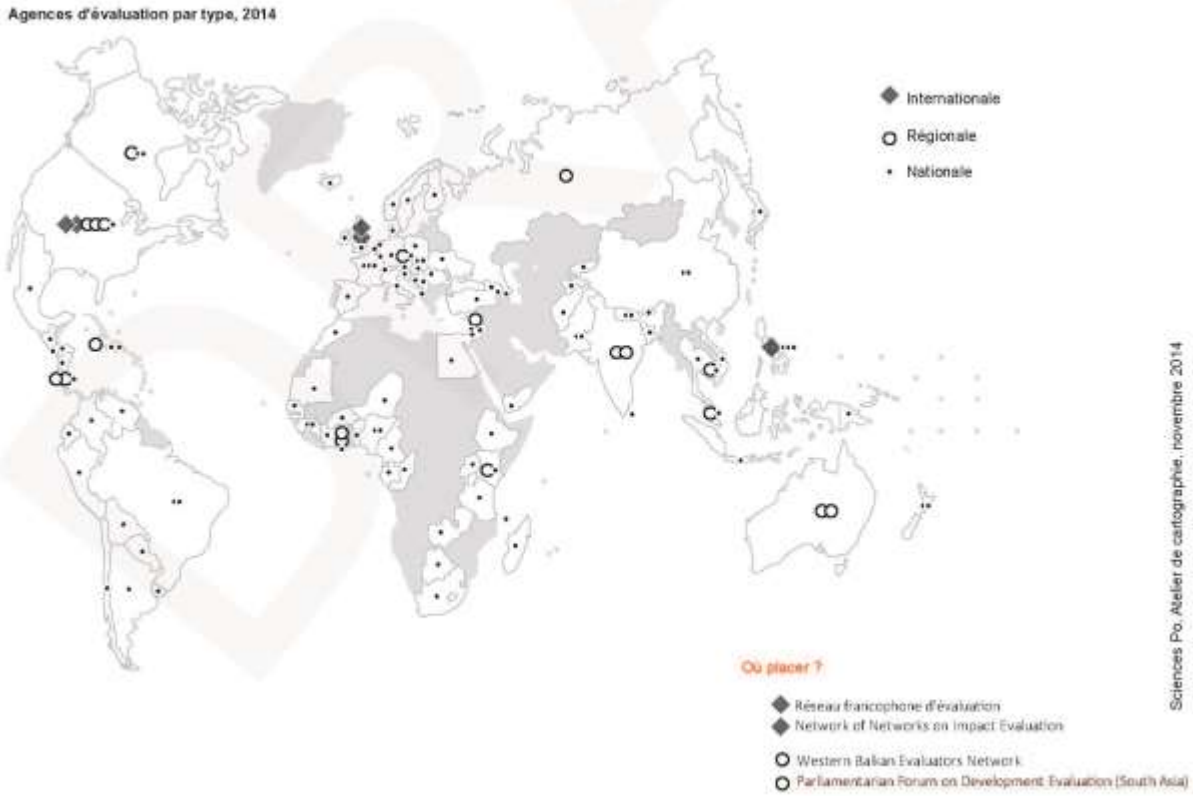
Auteurs	Résumé de l'évaluation de l'efficacité de l'aide
Griffin et Enos (1970), Weisskopf (1972)	Les flux d'aide font baisser l'épargne dans les pays bénéficiaires du même montant. La fongibilité de l'aide fait que l'activité marginale entrainée par l'aide ne mène pas à une hausse de l'accumulation de capital.
Mosley (1985)	"Micro-macro gaps": il y a un contraste entre l'inefficacité de l'aide au niveau internationale et son efficacité au niveau local
Cassen et al (1994)	50% des projets de développement sont efficaces et seulement très peu aggravent la situation des pays bénéficiaires, même s'ils échouent à atteindre leurs objectifs.
Hadji-Michael et al (1995)	L'aide a un effet positif sur la croissance mais la performance dépend de leurs politiques économiques (investissement public, déficit budgétaire, inflation).
Boone (1996)	L'aide a un effet positif, mais statistiquement pas significatif sur la croissance. La totalité de l'aide privée et publique est consommée par les Etats. L'aide entraîne donc seulement une hausse de la consommation publique. Or, il est reconnu que la hausse de la consommation publique a un effet négatif sur la croissance.
Burnside et Dollar (1997)	L'aide a un impact positif sur la croissance si la politique macro-économique de l'État bénéficiaire est saine. L'aide n'a pas la capacité d'influencer cette politique.
Tsikata (1998)	L'aide n'a pas d'impact significatif sur la croissance des pays. Il n'y a d'impact positif sur l'épargne et l'investissement que s'il y a effort d'ajustement. La qualité des instruments de politique macro-économique est déterminante pour la compétitivité et pour freiner l'éviction de l'investissement privé.
Durbarry, Gemmel et Greenaway (1998)	L'aide a un effet positif sur la croissance dans un environnement macro-économique stable.
Guillaumont et Chauvet (1999)	Une bonne politique macro-économique a une influence positive sur la croissance indépendamment de la présence d'aide ou de l'environnement extérieur.
Lensink et Morrissey (1999)	L'impact de l'aide sur la croissance est fonction des effets de l'aide sur l'investissement. La corrélation est systématiquement négative entre le degré d'incertitude quant aux apports d'aide et la hausse de la croissance économique.
Svenson (1999)	L'aide n'est efficace que dans les démocraties.
Collier et Dollar (1999)	L'allocation actuelle de l'aide est très loin de du ratio efficience/pauvreté optimal qu'il faudrait atteindre pour une meilleure efficacité de l'aide.
Hansen et Tarp (1999)	L'aide augmente l'épargne globale et l'investissement et a donc un effet positif sur la croissance. L'aide est utile, même si les Etats n'adoptent pas de bonnes politiques.
Guillaumont et Chauvet (2004) Rodrik (1996)	L'aide a tendance à perpétuer la mauvaise gouvernance de par son effet désincitateur sur la mise en œuvre des réformes.

Easterly et al (2003)	L'aide est systématiquement inefficace. Elle doit être abandonnée pour mieux laisser les pauvres prendre leur destin en main car ils sont les mieux informés et outillés pour le faire
Clemens, Radelet, Bhavnani (2004)	Catégorisation de l'efficacité de l'aide selon les types d'aide: l'aide à la construction d'infrastructure et l'aide budgétaire ont des effets positifs rapides sur la croissance. L'aide vers les secteurs sociaux a un effet de long terme et indirect. L'aide humanitaire n'a pas d'effet sur la croissance. Paradoxe "micro-macro": malgré des résultats décevants au niveau macro, il y a des preuves d'interventions d'aide ciblées réussies au niveau micro. L'aide sectorielle est efficace spécifiquement au niveau micro.
Rajan et Subramanian (2005)	L'aide a un impact négatif systématique sur la compétitivité des pays. Cela est dû à la surévaluation des taux de change engendrée par les flux d'aide.
Sachs (2006)	L'aide est efficace et ses montants doivent être augmentés afin de générer un "big push" sur le développement des pays bénéficiaires. La mauvaise gouvernance fait partie de la "trappe à pauvreté". L'aide aura un impact plus fort dans les États qui offrent de bonnes perspectives d'amélioration de la gouvernance, quel qu'en soit le niveau au départ.
Collier (2006)	En l'absence d'aide au développement, la croissance des pays africains aurait été pire qu'elle ne l'est aujourd'hui.
Birdsall (2007)	Une des causes de la performance limitée de l'aide est le faible investissement dans le renforcement des institutions et donc l'incapacité à résoudre "la trappe institutionnelle" qui empêche la croissance économique des pays du Sud.
Eifert, Gelb (2005)	La volatilité de l'aide nuit significativement à son efficacité.
Banerjee et al (2007)/Duflo Esther (2009)	L'aide peut être efficace mais le manque d'évaluations pour déterminer quels programmes fonctionnent réellement cause des gaspillages et des inefficiences considérables. Cela entraîne un pessimisme infondé sur le rôle de l'aide dans l'impulsion du développement économique.
Chataîgner, Magro (2007)	L'assistance financière dans les quatre années qui suivent le déclenchement d'une réforme diminue les chances qu'elle soit soutenue. L'assistance technique est à privilégier. L'influence devient visible cinq à huit années après la fin de l'intervention.
Bourguignon, Sundberg (2007)	L'efficacité de l'aide dépend de l'appropriation par les pays bénéficiaires et de son allocation sélective, basée sur la performance des programmes d'aide et des pays bénéficiaires.
Riddell (2007)	Le soutien de l'opinion publique à l'aide est indépendant de son efficacité et reste élevé. Les programmes d'aide atteignent généralement leurs objectifs directs, et le ratio de réussite a tendance à augmenter, mais l'impact de long terme est plus difficile à mesurer. Les évaluations "macro" régressives internationales ne prouvent rien à l'efficacité de l'aide. L'efficacité est dépendante de l'appropriation et de l'engagement des bénéficiaires.
Roodman (2007)	L'efficacité de l'aide existe à des niveaux multiples, et c'est l'agrégation des niveaux intermédiaires d'efficacité qui déterminent l'efficacité au niveau global.
Chauvet, Cogneau, Naudet (2008)	L'efficacité de l'aide dépend du niveau de "handicap structurel" de chaque pays.
Easterly (2009)	Une aide bien gérée produit des résultats positifs au niveau local dans des domaines tels que la santé ou l'éducation. Il faut favoriser les approches "marginales" sous la forme d'interventions ciblées de petite taille.
Ndikumana (2012)	La systématisation des évaluations des projets et programmes d'aide peut permettre de réduire le "micro-macro gap" sur l'efficacité de l'aide.

Source : GABAS, Jean-Jacques, LAPORTE, Camille, 2011

Annexe 4 : Émergence des réseaux d'évaluateurs





Annexe 5 : Excuses pour éviter l'évaluation

Ce tableau présente une sélection d'excuses pour éviter la mise en place d'évaluations rassemblées par un évaluateur des Nations unies :

1. Notre projet est spécifique	23. Le projet n'est pas évaluable
2. Cela va couter trop cher	24. La direction veut juste nous sanctionner
3. Nous n'avons pas le temps	25. L'évaluation va perturber nos relations avec les partenaires locaux
4. Le projet est trop petit	26. L'évaluation ne peut pas être objective
5. Ce n'était pas prévu dans la planification du projet	27. C'est trop difficile de changer les choses
6. C'est un projet expérimental	28. On a toujours travaillé comme cela
7. Le gouvernement ne va pas aimer cela	29. Nous avons fait ce que nous avions dit que nous ferions
8. Donner moi l'argent que vous voulez dépenser pour l'évaluation	30. Nous avons déjà été évalué
9. Les personnes externes au projet n'en comprendront pas la spécificité	31. Nous n'avons aucun problème avec ce projet
10. L'évaluation est une « tour d'ivoire »	32. Il y a eu un changement dans le gouvernement
11. Je pars en vacances	33. La crise financière a retardé le projet
12. Ce n'est pas notre préoccupation	34. Nous venons juste d'avoir été audités
13. Pourquoi changer le projet ? Il fonctionne très bien	35. C'est un de nos meilleurs projets
14. Nous ne sommes pas encore prêts	36. Le projet est trop récent/ trop ancien
15. Ce n'est pas dans le budget	37. Le projet n'est pas encore tout à fait terminé
16. Notre interlocuteur local sur ce projet est parti	38. Nous ne trouvons pas le document de planification originel du projet
17. Notre interlocuteur local sur ce projet a changé récemment	39. Je n'étais pas le responsable du projet lorsqu'il a débuté
18. Le directeur du projet n'a pas encore été nommé	40. Le gouvernement est satisfait du projet
19. Les opérateurs du projet sont en formation	41. Le gouvernement n'a pas encore fourni ses inputs
20. Nous nous débrouillons très bien sans évaluation	42. Nous n'avons pas les données
21. Je n'ai pas besoin que quelqu'un vienne m'apprendre mon métier	43. Le cadre logique du projet est trop vague
22. Nous nous autoévaluons continuellement	44. C'est la saison des pluies

Source : UNICEF, IPEN, *New Trends in Development Evaluation*, Evaluation working papers, issue 5, 2006

Annexe 6 : Logiques d'allocation de l'aide des donateurs

	2005	2006	2007	2008	2009	2010	2011
CPR Banque mondiale	0,47	0,45	0,467	0,541	0,483	0,539	0,458
Portfolio rating Banque mondiale						0,153	0,158
Gouv éco	0,334	0,482	0,431	0,398	0,317	0,491	0,438
Gouv institutionnelle	0,551	0,518	0,601	0,56	0,661	0,591	0,627
NEI DFID							0,558
avant/après CPIA				0,972	0,912	0,923	
NEI/CPIA							0,006
NEI/fragilité							0,56
AUSAID perfomance/allocation							0,177
USAID/World Gouvernance Index					0,298	0,267	0,284
USAID/OMP et actions militaires							0,79

Annexe 7 : Lexique

Aide budgétaire : Aide destinée à financer le budget de l'État sans affectation à des dépenses pré-identifiées ; elle est inscrite en ressources du budget de l'État et décaissable selon certains critères prédéfinis : indicateurs de mise en œuvre des CSLP, d'amélioration de la gestion des finances publique.

Aide liée/déliée : L'aide est "liée" lorsqu'elle permet de financer l'acquisition de biens et de services exclusivement auprès de fournisseurs appartenant au pays qui a apporté les fonds d'aide. Elle est "déliée" lorsqu'elle peut être utilisée pour acheter ces biens et services auprès de n'importe quel pays partenaire.

Alignement : Processus par lequel les bailleurs de fonds prennent comme cadre de référence les politiques nationales sectorielles ou macro-économiques définies par les pays partenaires, notamment dans les CSLP. Ce terme est utilisé depuis la Conférence internationale de Monterrey en mars 2002

Audit : L'audit sert à contrôler la légalité la conformité et la régularité de l'utilisation de ses ressources par une organisation. Il est réalisé de manière périodique selon des normes préétablies dans les statuts légaux de l'organisation. L'audit sert à déceler et prévenir des irrégularités dans la gestion organisationnelle et financière d'une organisation. L'audit peut être réalisé par des auditeurs externes ou internes, ou encore par la Cour des comptes.

Bonne gouvernance : La notion de « bonne gouvernance » incarne l'idée d'un gouvernement pragmatique, orienté vers l'action mais aussi plus disposé à négocier les choix collectifs rationnels en consultation avec les acteurs sociaux concernés

Bonnes pratiques : Les « bonnes pratiques » sont un ensemble de règles, instructions ou codes de conduite fondés sur l'idée qu'il existe, pour atteindre un résultat donné, une méthode meilleure que les autres. Issues de l'expérience passée au filtre de l'étude statistique, les bonnes pratiques sont donc à la fois un standard qui prétend décrire objectivement les voies de la performance, en même temps qu'un guide pour l'action. En somme, les bonnes pratiques ont pour vocation de normaliser les comportements, afin de réduire l'écart entre le fait et le prescrit.

Cadre logique : Le cadre logique est une méthode de gestion axée sur les résultats (GAR). Il permet de vérifier la cohérence entre les différents niveaux de projets, en prenant comme point de référence son objectif le plus élevé. Développé dès la fin des années 60 par la coopération américaine (USAID), le cadre logique a été repris depuis lors par la quasi-totalité des

organisations de développement. Le cadre logique synthétise sous forme de tableau toutes les informations clés d'un projet: objectifs, résultats, activités, risques, programmation, ressources. La méthode du cadre logique est la succession d'étapes et d'analyses qui va permettre l'élaboration progressive de la matrice.

Conditionnalité : Lorsqu'un pays emprunte auprès d'une institution financière, son gouvernement prend des engagements. Ce sont les conditionnalités. Elles sont destinées à permettre de vérifier que l'argent prêté est bien utilisé pour résoudre les problèmes qui ont motivé le prêt ou le don. En général, elles accompagnent l'échelonnement des décaissements, chacun d'entre eux permettant de vérifier que les engagements sont respectés.

Contrôle : Le contrôle est une activité réalisée en interne d'une organisation. Il désigne l'ensemble des dispositifs (processus), organisés, formalisés et permanents, choisis par l'encadrement de cette organisation, mis en œuvre par les responsables de tous niveaux pour maîtriser le fonctionnement de leurs activités; ces dispositifs sont destinés à fournir une assurance raisonnable quant à la réalisation de leurs objectifs : 1) la réalisation et l'optimisation des opérations ; 2) la protection des actifs et des ressources financières ; 3) la fiabilité des informations comptables et financières ; 4) la conformité des lois et des règlements ; 5) la qualité des prestations

Empowerment : L'empowerment est l'octroi de plus de pouvoir aux individus ou aux groupes pour agir sur les conditions sociales, économiques, politiques ou écologiques qu'ils subissent. Diverses traductions ont été proposées en français : « capacitation », « développement du pouvoir d'agir », « autonomisation », « responsabilisation », « émancipation » ou « pouvoir-faire ». Ce terme est flou et son sens varie selon les acteurs qui l'emploient. Les institutions internationales comme la Banque mondiale ont commencé à utiliser la notion d'empowerment dans les politiques de développement, notamment en direction des femmes. Toutefois, elle reflète souvent une vision individualiste et néolibérale, où les gens se prendraient en charge pour pallier les manquements des États dans le champ des droits fondamentaux.

Évaluation : Appréciation systématique et objective d'un projet, d'un programme ou d'une politique, en cours ou terminé, de sa conception, de sa mise en œuvre et de ses résultats. Le but est de déterminer la pertinence et l'accomplissement des objectifs, l'efficience en matière de développement, l'efficacité, l'impact et la durabilité. Une évaluation devrait fournir des informations crédibles et utiles permettant d'intégrer les leçons de l'expérience dans le processus de décision des bénéficiaires et des bailleurs de fonds. Le terme "évaluation" désigne également

un processus aussi systématique et objectif que possible par lequel on détermine la valeur et la portée d'une action de développement projetée, en cours ou achevée.

Évaluation expérimentale : L'évaluation expérimentale aléatoire contrôlée, dans sa version la plus élémentaire, implique de constituer deux groupes de personnes, puis de donner l'accès au dispositif que l'on souhaite évaluer à un groupe que l'on nommera « le groupe test » et de ne pas donner l'accès à l'autre groupe que l'on nommera « le groupe témoin ». Durant la mise en œuvre du dispositif, l'évolution des conditions de vie des deux groupes seront suivie. Une fois le dispositif achevé, les écarts observés entre les deux groupes permettent d'évaluer les « impacts » du dispositif sur le groupe test.

Évaluation participative : L'évaluation participative est conçue afin d'aider l'individu à s'auto-assister. Il s'agit d'un processus démocratique dans lequel les individus renforcent leurs propres pouvoirs avec l'aide d'un expert extérieur qui fait principalement office d'encadreur et d'animateur. Les participants au programme décident de procéder à leur propre évaluation, ils la planifient, ils la mettent en œuvre, ils collectent et analysent leurs propres données, ils interprètent leurs résultats et tirent leurs propres recommandations et enseignements. Enfin, ils mettent en œuvre leurs propres recommandations. Le rôle de l'évaluateur professionnel externe est ici fort différent de celui de l'évaluateur dans le cas des évaluations traditionnelles. Il/ elle doit travailler directement avec les parties prenantes, pour qu'ils effectuent leur évaluation, et non pas travailler pour eux, pour effectuer son évaluation à lui/ elle. L'évaluateur externe doit jouer un rôle d'encadreur, animateur, formateur, conseiller et non pas celui de l'évaluateur classique.

Evidence based policy making : Evidence based policy making réfère à une politique publique mise en œuvre sur la base de preuves d'efficacité rigoureusement et objectivement établies. Ces politiques sont établies après la mise en place d'études scientifiques et rigoureuses, telles que les évaluations aléatoires contrôlées identifiant les programmes et pratiques capables d'améliorer les résultats des politiques publiques.

Fonds verticaux : Initiatives internationales en dehors du système de l'ONU, levant des fonds significatifs au niveau des pays pour soutenir des objectifs thématiques ciblés, c'est-à-dire l'atteinte des objectifs du Millénaire pour le Développement.

Gestion axée sur les résultats : La gestion axée sur les résultats vise à fournir un cadre cohérent de planification stratégique et de gestion fondé sur l'apprentissage et la responsabilité dans un environnement décentralisé. C'est en premier lieu un système de gestion, et en second lieu un

système d'information sur les résultats. L'adoption d'une telle approche vise à accroître l'efficacité de la gestion et la responsabilisation des gestionnaires par la définition de résultats réalistes à atteindre, le suivi des progrès vers ces résultats, l'utilisation des leçons apprises dans les processus de décisions et d'établissement de rapports sur les résultats finaux.

Groupe de référence : Comité constitué de parties prenantes de la politique publique évaluée et d'experts extérieurs compétents chargé de suivre la bonne mise en œuvre de l'évaluation et d'en valider les conclusions et les recommandations.

Impact : Effets à long terme, positifs et négatifs, primaires et secondaires, induits par une action de développement, directement ou non, intentionnellement ou non. Effets d'une action affectant, plus ou moins, la capacité d'un pays ou d'une région d'utiliser ses ressources propres (humaines, financières et naturelles), de façon plus efficiente, équitable et durable.

Réalisation : Activités mises en œuvre dans le cadre d'un projet ou programme d'aide au développement

Redevabilité : Obligation de rendre compte du fait que le travail a été conduit selon les règles et les normes convenues, ou obligation de rendre compte de façon claire et impartiale sur les résultats et la performance, au regard du mandat et/ou des objectifs fixés. Cela peut exiger une démonstration précise, voire juridique, que le travail a été accompli selon les termes du contrat.

Remarque : dans le contexte du développement, le terme peut concerner l'obligation des partenaires à agir selon des responsabilités clairement établies, un partage des rôles et des performances attendues, souvent dans le respect d'une gestion prudente des ressources. En ce qui concerne les évaluateurs, le terme évoque la responsabilité de fournir des appréciations de performance et des rapports d'avancement précis, impartiaux et crédibles. Pour les décideurs et les gestionnaires du secteur public le devoir de rendre compte s'effectue envers les contribuables et les citoyens.

Redevabilité mutuelle : Les donateurs et leurs pays partenaires sont réciproquement redevables des résultats de développement apportés par les projets et programmes d'aide au développement.

Résultat : Effet direct et immédiat d'un projet ou programme d'aide au développement sur les bénéficiaires qu'elle cible

Suivi : Processus continu d'examen de l'apport des réalisations d'une intervention à ses bénéficiaires escomptés, qui est mené durant l'exécution de ladite intervention dans l'intention

de corriger immédiatement tout écart par rapport aux objectifs opérationnels. Le suivi est un système continu de production et de collecte de données et d'informations relatif à l'exécution d'une intervention au regard des objectifs opérationnels et des échéances convenus. Il permet de fournir en permanence aux responsables et aux parties prenantes des informations en retour sur l'exécution de l'intervention et de procéder aux ajustements nécessaires en temps opportun. Le suivi permet d'ajuster régulièrement le contenu et les méthodes de l'intervention sans toutefois remettre en cause les grandes options ni réorienter l'action en profondeur.

Système d'évaluation : Dans une organisation, on parle de système d'évaluation mature lorsqu'il y a une demande d'évaluation durable à des niveaux et dans des secteurs variés à laquelle répond une offre de services en évaluation compétents, que les résultats des évaluations sont régulièrement utilisés dans le secteur public et qu'un cadre institutionnel assure la confiance en la qualité des évaluations. Le terme système suggère que les quatre composantes citées sont intégrées dans un réseau de cercles vertueux qui permettent à une culture de l'évaluation d'émerger et de se maintenir. Le système inclue des acteurs individuels, des organisations, des réseaux et des mécanismes

Source : AFD, Dictionnaire du Développement : http://www.afd.fr/home/AFD/dictionnaire-developpement; CAD-OCDE, *Glossaire du CAD de l'OCDE sur les principaux termes relatifs à l'évaluation*, 1991 ; auteure

Annexe 8 : Liste des entretiens menés

1. Ancien directeur du département de l'évaluation de l'AFD, 21 décembre 2011
2. Deputy Head of Programme - EU external action programme at ECDPM and Visiting Professor at College of Europe de Bruges, 31 janvier 2011
3. Chargée d'études en évaluation, UNESCO, 6 avril 2012
4. Chargée de projets, F3E, 19 avril 2012
5. Évaluateur au sein de la JICA, entretien Skype, 8 mai 2012.
6. Ex présidente de l'AFREA, 22 mai 2012
7. Responsable pays, agence de la Banque mondiale au Laos, 12 juin 2012
8. Consultante pour le PNUD au Laos, 12 juin 2012, 10 juillet 2012
9. Représentant du PNUD au Laos, 13 juin 2012
10. Responsable communication externe, agence de la Banque mondiale au Laos, 13 juin 2012
11. Directeur des agences du Laos et du Cambodge de l'AFD, 19 juin 2013
12. Chargé de projets de l'AFD au Laos, 30 juin 2012
13. Chargé de projets pour l'aide au développement chinoise au Laos, 18 juillet 2012
14. Membre de l'agence de la JICA au Laos, Vientiane, 25 juin 2012.
15. Membre du réseau d'évaluateurs du CAD-OCDE, 16 novembre 2012
16. Consultant indépendant en évaluation, 12 janvier 2013
17. Ancien évaluateur interne de l'ODA, 14 janvier 2013.
18. Évaluateur retraité du ministère de la coopération français, consultant auprès d'organisations internationales, 21 mars 2013

19. Évaluateur du ministère de la coopération français, consultant auprès d'organisations internationales, 4 avril 2013

20. Directeur d'un cabinet d'évaluation français, 12 avril 2013

    Executive Director, National Center for Science and Technology Evaluation (NCSTE) en Chine, 8 juin 2013

21. Évaluateur à l'AFD, 14 juin 2013

22. Ancien évaluateur au DFID, 17 juin 2013

23. Évaluatrice au DFID, 12 juillet 2013

24. Ancien directeur du département de l'évaluation de l'AFD, 17 septembre 2013

25. Évaluatrice à l'AFD, 8 octobre 2013

26. Évaluatrice à la Banque mondiale, 24 octobre 2013

27. Évaluateur à l'AFD spécialiste des évaluations d'impact, 19 novembre 2013

28. Directeur d'un cabinet d'évaluation, consultant en évaluation auprès du CAD-OCDE, 20 novembre 2013

29. Évaluatrice retraitée de la Banque mondiale, 20 novembre 2013

30. Directrice de l'Observatoire du Changement Social du Bénin, 3 décembre 2013

31. Directeur de division à l'AFD, 5 décembre 2013

32. Évaluatrice à la Banque mondiale, entretien à Paris, le 20 décembre 2013.

33. Directeur d'un cabinet d'évaluation des politiques publiques, 13 janvier 2014

34. Co-responsable du réseau d'évaluateurs IOCE, 22 janvier 2014

35. Ancienne Chief Social Development Advisor au DFID, consultante pour des agences de développement, 14 février 2014

36. Consultant en évaluation auprès d'organismes d'aide au développement, chercheur en économie, 18 février 2014.

37. Directeur d'un cabinet d'évaluation au Sénégal, président de la société sénégalaise d'évaluation, 3 avril 2014

38. Chargé de projet pour une agence d'aide au développement au Sénégal, 3 avril 2014

39. Agent du ministère de l'environnement et du développement durable du Sénégal, 5 avril 2014

40. Représentant de la Commission européenne au Sénégal, entretien Skype, 7 avril 2014

41. Évaluateur à la Banque mondiale, 21 mai 2014

42. Évaluateur à la KOICA, entretien à Seoul, le 16 juin 2014

43. Chargé de projets sur l'efficacité de l'aide à la KOICA, entretien à Séoul, le 16 juin 2014

44. Ancien agent de la KOICA, entretien à Séoul, 17 juin 2014

45. Ancien agent du DFID, agent du ministère des affaires étrangères, 19 juin 2014

46. Chercheur économiste à l'AFD, 23 juin 2014

47. Évaluatrice à l'AFD, 25 juin 2014

48. Senior Policy Advisor au sein du département de l'efficacité de l'aide du CAD-OCDE, 2 juillet 2014

49. Analyste politique, unité d'évaluation du CAD-OCDE, 2 juillet 2014

50. Analyste politique, département de l'efficacité de l'aide, CAD-OCDE, 2 juillet 2014

51. Senior evaluation advisor, DFID, 4 juillet 2014

52. Consultant en évaluation des politiques publiques au Maroc, 17 septembre 2014

53. Évaluateur au sein des Nations unies, coprésident de l'initiative IOCE 6 octobre 2014

54. Évaluateur à la JICA, entretien à l'AFD, 28 novembre 2014

55. Directeur d'un cabinet d'évaluation des politiques publiques, 9 décembre 2014

Annexe 9 : Liste non exhaustive des évènements observés

Agence française de Développement

Réunions mensuelles de coordination de la division de l'évaluation de l'AFD

Comité de pilotage du barrage X au Laos, 22 octobre 2012, 16 janvier 2013

Laurent Fontaine, directeur de la division de l'évaluation de l'AFD

Yves Ficatier, évaluateur à l'AFD

André Pouilles-Duplaix, directeur des agences de l'AFD au Laos et au Cambodge

Dominique Richard, chef de projet énergie à l'AFD

Alain Ries, directeur de l'unité transports et énergie durable

Catherine Garreta, directrice du département des appuis transversaux

Jean-Noel Roulleau, appui environnemental et social, AFD

Jean-Yves Grosclaude, directeur exécutif en charge des opérations

Nicolas Fornage, directeur de l'agence de l'AFD au Pakistan, anciennement chargé de projet de grandes infrastructures

Grégory Clémente, directeur du département des risques, AFD

Réunions de préparation de la politique d'évaluation de l'AFD janvier-mars 2013

Réunion des équipes d'évaluation de l'AFD, du ministère des affaires étrangères et de la DG Trésor, 19 décembre 2013

Équipe d'évaluation de l'AFD

Équipe d'évaluation de l'aide au développement du Ministère des affaires étrangères

Équipe d'évaluation de l'aide au développement de la DG Trésor

Benoit Lajudie, commissariat général à la stratégie et à la prospective, ancien responsable de l'évaluation à la DGOM

Fabrice Alcandre, chargé de projet, cabinet d'évaluation des politiques publiques C2G

Luc Lefebvre, directeur du cabinet d'évaluation des politiques publiques S3E

Table ronde sur la coopération avec les donateurs émergents avec le DFID et la JICA, 28 novembre 2014

Philippe Orliange, directeur du département de la stratégie, des partenariats et de la communication, AFD

Représentants du département de la Recherche de l'AFD

Représentants des départements opérationnels et stratégiques de l'AFD

Frédéric Bontemps, directeur de la direction des biens publics mondiaux, ministère des affaires étrangères

Représentants de la JICA

Représentants du DFID

Groupes de travail de réseaux d'évaluateurs

Réunion du groupe de travail de la SFE sur l'évaluation de l'aide au développement, « Quelle utilisation et quelle influence de l'évaluation », 15 novembre 2012

Olivier Picavet, chargé de projets au Secrétariat général de la Modernisation de l'Action publique

Charlotte Boisteau, chargé de projets au F3E

Laurent Fontaine, directeur de la division de l'évaluation de l'AFD

Hans Lundgren, président du réseau d'évaluateurs du CAD-OCDE

Michael Ruleta, consultant en évaluation

Anne-Gaëlle Nuts, consultante en évaluation

Marc Raffinot, professeur en économie à l'université Paris Dauphine

Marion Lory, organisation internationale de la francophonie

Ximena Rodriguez, cabinet d'évaluation Euréval

Christian Hugues, directeur du cabinet d'évaluation Evalua

Blaise Gonda, chargé de projets au sein du cabinet d'évaluation Louis Berger

Jean-François Sempère, économiste au sein du cabinet Louis Berger

Journée de travail SFE « évaluation du développement : quelle perspective ? », 28 février 2013

Oumoul Ba Tall, directrice de la société de conseil OKT-Consult en Mauritanie

Hubert de Milly, ancien évaluateur au ministère de la coopération français, consultant auprès d'organisations internationales

Christine Andela, member des ONG COSADER et National Alliance against Hunger, Cameroun

Grégoire Laourou, président de la Commission des Finances, Assemblée nationale, Bénin

Jean-Louis Viélajus, président de Coordination Sud

Sébastien Mosneron-Dupin, administrateur principal au sein de la Commission des affaires étrangères au Sénat

Nermine Wally, présidente de l'Association africaine d'évaluation

Dorte Kabell, consultante en évaluation, Danemark

Justine A Odjoubé, directrice de l'Observatoire du changement social du Bénin, ministère d'État chargé de la prospective, du développement et de l'évaluation de l'action publique

Mayacine Camara, responsable de l'unité de coordination et de suivi de la politique économique et de suivi de la politique économique du ministère de l'économie et des finances du Sénégal

Mamadou A Dembele, directeur du secrétariat à l'harmonisation de l'aide au sein du ministère de l'économie et des finances du Mali

Laurent Fontaine, directeur de la division de l'évaluation de l'AFD

El Hadji Gueye, coordonnateur du Centre CLEAR de l'Afrique francophone

Boureima Gado, président du réseau nigérien de suivi et d'évaluation

Journées françaises de l'évaluation, « l'Évaluation des politiques publiques, État de l'art et controverses », 6-7 juin 2013

Groupe de travail Evaled, « le Renforcement des capacités en évaluation », 17 octobre 2013

Charlotte Boisteau, chargée de projets au sein du F3E

Nathalie Garcia, évaluatrice au Secours Catholique Caritas France

Jean-Marie Loncle, président du réseau francophone de l'évaluation

Marion Lory, chargée des questions de l'évaluation au sein de l'organisation internationale de la Francophonie

Sébastien Galéa, directeur de l'association EVAL

Bruno de Reviers, chargé de projets au sein du F3E

Groupe de travail SFE-EVALED, « associer des professionnels locaux dans les évaluations des interventions d'appui au développement – un plus ou une obligation ? », jeudi 30 janvier 2014

Véronique de Rohan-Chabot, directrice de la division de l'évaluation, ministère des affaires étrangères

Laurent Fontaine, directeur de la division de l'évaluation de l'AFD

Fabrice Alcandre, directeur du cabinet d'évaluation, C2G conseil, France

Arnauld Bertrand, consultant en évaluation, Ernst and Young, France

Vincent Coppens, consultant en évaluation, ADE, Belgique

Luc Lefebvre, directeur du cabinet d'évaluation SEE, Luxembourg

Jean-Claude Le Goff, SERES, Belgique

Ahmed Bencheikh, Consultant à Evalua Development & Evalua Institute, Maroc

Boureima Gado, président du réseau nigérien de suivi et d'évaluation, Niger

Moctar Saw, consultant en évaluation, président de la société sénégalaise d'évaluation, Sénégal

Oumoul Khayri Ba Tall, directrice de la société OKT-Consult, Mauritanie

Au Laos (juin-juillet 2012)

Visite du site du barrage X et rencontre avec les opérateurs du projet, 5-7 juin 2012

Visite des villages de populations déplacées par le barrage X, 25-26 juin 2012

Échanges réguliers avec le personnel de l'entreprise opérateur du projet de barrage X à Vientiane, juin-juillet 2012

Réunions de suivi-évaluation des donateurs des compensations environnementales et sociales du barrage X

Représentants de l'AFD

Représentants de la Banque mondiale

Représentants du PNUD

Représentants de l'opérateur du projet

Réunion de coordination de l'évaluation des effets environnementaux et sociaux du barrage X, 17 juillet 2012

Keiko Miwa, responsable pays, agence de la Banque mondiale au Laos

Elizabeth Mann, consultante pour le PNUD au Laos

Bui Duy Thanah, représentant du PNUD au Laos

Martin Gray, responsable communication externe, agence de la Banque mondiale au Laos

Ruedi Luthi, représentant de l'opérateur du projet de barrage

André Pouillès-Duplaix, directeur des agences de l'AFD au Laos et au Cambodge

Khonesavahn, Xaymoungkhoune, responsable de projets, agence de l'AFD au Laos

Évaluations observées

Évaluation des effets environnementaux et sociaux du barrage X au Laos (février 2012- mai 2013)

Évaluation du réseau culturel français à l'étranger, Cour des comptes (mission de terrain au Maroc) janvier-juin 2013

Évaluation de l'Initiative nationale de Développement humain au Maroc, KPMG Maroc

Évaluation de la coopération française dans le domaine de la santé, ministère des affaires étrangères (mission de terrain en Côte d'Ivoire)

Ces entretiens ont été complétés par de nombreuses discussions informelles avec des évaluateurs et « évalués » lors de missions d'expertise à l'AFD, au ministère des Affaires étrangères (Au Siège à Paris, en Suisse et en Côte d'Ivoire), à la Cour des Compte (À Paris et lors d'une mission d'évaluation au Maroc sur le réseau culturel français), ainsi qu'au sein de cabinets privés de consultance en évaluation.

Annexe 10 : Profils d'évaluateurs

Évaluateurs externes aux organismes d'aide

Éric Monnier, co-directeur du cabinet d'évaluation des politiques publiques Euréval

Éric Monnier est Ingénieur du Génie Civil de l'Institut National des Sciences Appliquées de Toulouse (INSA) en 1976 et Docteur en Sociologie urbaine et Aménagement du territoire en 1982. Il a débuté sa carrière professionnelle au Centre Scientifique et Technique du Bâtiment, où il découvre l'évaluation grâce à son chef de service, Michel Conan, qui s'intéressait lui-même beaucoup à l'évaluation. Il est ensuite parti aux États-Unis pour se former à l'évaluation et travaille aux Nations-Unies. Il est recruté en 1989 à l'École Nationale des Travaux Publics de l'État (ENTPE) pour mettre en place un centre de recherche sur l'évaluation. Il créera ainsi le laboratoire CEOPS (Conception d'Évaluation pour les Organisations et les Politiques publiques).

Occupant de nombreuses positions institutionnelles dans le champ de l'évaluation, il est devenu membre du Conseil Scientifique de l'Évaluation en 1992, où il a contribué à l'élaboration d'une « doctrine à la française ».

En 1992, il a entamé une carrière dans le privé en fondant son cabinet de conseil dédié à l'évaluation des politiques publiques, le cabinet C3E (qui devient au début des années 2000, Euréval) qui est à l'origine de nombreux guides méthodologiques en évaluation.

Justine Odjoube, évaluatrice béninoise

Justine Odjoube est diplômée de l'École Polytechnique du Bénin. Elle est spécialisée en statistiques et en économie appliquée.

En 1982, à la suite de ce diplôme, elle a commencé sa carrière professionnelle au ministère du plan, de la statistique et de l'analyse économique, plus précisément au Bureau Central des Projets. Elle était chargée de la réalisation d'études de préfaisabilité et de faisabilité des projets de développement. Elle appréciait beaucoup ce travail et a souhaité poursuivre ces études pour se spécialiser dans l'analyse économique des projets de développement.

Ainsi, elle a interrompu sa carrière en 1986 pour faire un Doctorat à l'Institut de Démographie de l'Université Catholique de Louvain-la-Neuve en Belgique. Après cette formation, elle a repris son travail au Ministère en charge du Plan, mais cette fois-ci à l'Institut National de la Statistique et de l'Analyse économique (INSAE), en qualité de chef du service des études démographiques, de chef du service Études et Perspectives de Population et de chef du service des statistiques agricoles. Durant les dix ans qu'elle a passés à l'INSAE, elle a dirigé plusieurs collectes et analyses de données, dont le deuxième et le troisième recensement général de la population et de l'habitation dans les départements de l'Atacora-Donga et du Plateau.

Pendant cette période, elle a pris part à plusieurs formations internationales portant sur l'analyse des données de recensement, la collecte, le traitement, l'analyse et l'utilisation des statistiques agricoles, et la confection et la gestion des bases de données.

Son implication effective dans la fonction de suivi a débuté au milieu des années 1990, lorsqu'elle était tant cadre de la direction des statistiques démographiques et Sociales au ministère de la planification du Bénin. Elle a commencé à pratiquer concrètement l'évaluation en 2004, après sa nomination à la tête de l'Observatoire du Changement Social (OCS), en qualité de Coordonnatrice. L'OCS est une structure du gouvernement béninois, chargé de suivre et d'évaluer l'impact de tous les programmes et projets mis en œuvre au Bénin dans le cadre de la réduction de la pauvreté. C'est aussi la structure en charge du suivi des Objectifs du Millénaire pour le Développement. Elle a dirigé l'OCS de mai 2004 à janvier 2013, date de son départ à retraite.

Christian Hugues, directeur du cabinet d'évaluation des politiques publiques Évalua

Au début des années 1990, Christian Hugues a fait des études de gestion à l'Université de Lyon avec la volonté de travailler dans le secteur public.

Il a exercé de 1990 à 1996 la responsabilité de contrôleur de gestion de la Ville de Blagnac (Haute-Garonne). Dans ce cadre, il a contribué à l'élaboration d'un système de management participatif par objectifs, au développement d'outils de contrôle de gestion, à l'émergence d'une fonction de marketing public et à la reconnaissance du caractère innovant des pratiques managériales de cette ville moyenne.

De 1994 à 1998, il a conduit des recherches théoriques sur le thème de l'évaluation des politiques publiques en conduisant parallèlement une recherche doctorale auprès de l'Université Paris IX-Dauphine, une activité de recherche et développement pour le compte du département secteur public-collectivités locales de KPMG-Fiduciaire de France et une recherche-action sur les besoins en formation des cadres territoriaux en charge des politiques contractuelles, pour le compte du Centre Supérieur de la Fonction Publique Territoriale.

En 1998, Christian Hugues a monté son cabinet d'évaluation des politiques publiques, Evalua, avec un associé spécialisé dans les méthodes quantitatives qu'il avait rencontré en Afrique quelques années auparavant. Son premier souhait a été de travailler dans le domaine de l'aide au développement. Il s'est donc rapproché du ministère des affaires étrangères, de l'AFD et de la DG Trésor et a rapidement réussi à collaborer avec ces structures. Aujourd'hui, le cabinet Evalua a une structure plus informelle, qui fonctionne comme un réseau d'évaluateurs.

Moctar Saw, évaluateur sénégalais

Après des études de sciences politiques et d'économie à l'Université de Dakar, Moctar Saw part suivre un Master de politiques publiques à Lyon au début des années 2000, où il se forme à l'évaluation des politiques publiques. À l'issue de ce Master, il fait un stage dans le cabinet

d'évaluation des politiques publiques, Evalua, à Toulouse. Passionné par l'évaluation, il souhaite développer cette pratique au Sénégal. Il monte alors son cabinet d'expertise en évaluation des politiques publiques et s'engage activement dans le renforcement des capacités d'évaluation au Sénégal. Il réussit à réunir un petit noyau de personnes intéressée par l'évaluation et créé l'association sénégalaise d'évaluation la Sénéval. Aujourd'hui, il participe à de nombreuses évaluations partenariales avec des donateurs et est un membre très actif des réseaux francophones sur l'évaluation.

Charlotte Boisteau, chargée de projet au F3E

En 2001, Charlotte Boisteau a été diplômée d'une double maîtrise en science politique et en sociologie à l'université de la Sorbonne Nouvelle. De 2011 à 2003, elle a suivi un DEA en étude du développement et relations internationales à l'Institut universitaire d'études du développement. Elle s'est ensuite lancée dans un doctorat en sociologie urbaine à l'École polytechnique fédérale de Lausanne.

De 2001 à 2005, elle a travaillé à l'UN-Habitat, tout d'abord en tant que coordinatrice de projet, puis de formatrice. Elle a ensuite rejoint le *United Nations Institute for Training and Research* où elle a assuré la fonction de coordinatrice de projet de 2004 à 2008. Durant cette période et jusqu'en 2009, elle a réalisé des missions de consultance en évaluation dans le secteur de l'aide publique en développement et de la recherche.

Elle a ensuite quitté les Nations unies pour devenir chargée des recherches et des enseignements pour le Samu social international. De 2009 à 2013, elle a assuré la fonction de chargée de travaux dirigés dans des universités et grandes écoles suisses, espagnoles et françaises. Depuis 2010, elle est chargée de mission « Europe et international » au sein de la F3E. Elle est très activement investie dans le groupe de travail sur l'évaluation de l'efficacité de l'aide au développement de la société française d'évaluation.

Driss Ennagadi, évaluateur marocain

Driss Ennagadi a une maîtrise en relations internationales. Il a travaillé au sein de différentes collectivités territoriales en France, notamment, à partir du milieu des années 2000 pour le département de l'Essonne. Il y travaille en tant que cadre territorial dans le département de l'audit interne. Dans le cadre de ces fonctions, il s'intéresse à l'évaluation. C'est alors qu'en 2007 la première division de l'évaluation a été créée au sein du conseil général de l'Essonne. Il témoigne de sa volonté d'intégrer cette division et suit une formation accélérée en évaluation. Il obtient le poste de de chef du service d'évaluation du conseil général. Son intérêt pour l'évaluation des politiques publiques ne fait que se renforcer au fil des années.

En 2011, il est sollicité, avec d'autres experts français, par le ministère de l'intérieur marocain pour former le secrétariat général du Maroc à l'évaluation des politiques publique, dans le cadre de la grande réforme de la gouvernance du pays. Prenant conscience de la demande d'évaluation

au Maroc et du besoin d'expert locaux, il monte son cabinet d'évaluation au Maroc et obtient rapidement, grâce à ses contacts, le marché de l'évaluation de l'initiative nationale de développement humain. Pour assurer la rentabilité et la pérennité de son cabinet, il en diversifie les missions aux études financières et au conseil politique.

Michele Tarsilla, consultant indépendant en évaluation auprès de la Banque mondiale

De 1995 à 2000, Michele Tarsilla suit une formation en affaires internationales et diplomatie à l'université catholique de Milan. Il part ensuite aux États-Unis pour suivre un Master en relations internationales et se spécialise dans la coopération en matière de santé. Pour financer ces études, il réalise des missions de consultance en tant qu'évaluateur pour les Nations unies, puis pour la Banque mondiale.

Il n'a pas de formation particulière en évaluation mais il aborde la théorie de cette activité dans le cadre de son Master à *Gorgetown University*. Il devient membre de la société américaine d'évaluation au milieu des années 2000. Il poursuit ses activités de consultant jusqu'en 2008, puis décide de suivre une formation qualifiante en évaluation afin de se professionnaliser et de gagner en légitimité. Il suit la formation de *l'International Program for Development Evaluation* (IPDET), puis commence un doctorat à la *Western Michigan University* qu'il achève en 2012. Durant cette période, il devient responsable du groupe de travail sur l'évaluation interculturelle au sein de la société américaine d'évaluation, poste qu'il occupe toujours aujourd'hui. À la suite de son Doctorat, il monte *l'Evaluation Capacity Development Group*.

Évaluateurs internes aux agences d'aide

Constance Corbier Barteaux, évaluatrice à l'AFD

Avant d'intégrer la division de l'évaluation de l'AFD, Constance Corbier-Barteaux était chef de projet au sein du département opérationnel « environnement et des ressources naturelles ». Elle a intégré la division de l'évaluation car elle avait envie de voir l'aide au développement sous un nouvel angle après plus de vingt ans à avoir travaillé dans l'opérationnel. Elle appréhendait cette prise de fonction, car elle connaissait mal le sujet de l'évaluation. Son profil n'était absolument pas spécialisé en évaluation. L'objectif qui lui avait été donné au sein de la division était de faire un état des lieux des projets de l'agence sur la biodiversité, afin de développer le portefeuille et la stratégie de l'AFD sur ce thème.

Nicolas Brillon, évaluateur à l'AFD

Nicolas Brillion a rejoint la division de l'évaluation en septembre 2010. Entré à l'AFD au service documentation, il s'est ensuite orienté vers l'opérationnel en tant que coordinateur géographique sur le Bénin et le Togo au sein du département Afrique. Il a ensuite occupé un

poste d'interface entre la Stratégie et les Opérations en représentant la Stratégie aux comités de projets, de l'identification à l'instruction. Il a rejoint la division de l'évaluation car il était attiré par le fait d'être à l'interface des différents niveaux de la hiérarchie de l'agence : les opérationnels, les divisions stratégiques et la direction. Toutefois, il se sentait peu légitime, car il n'avait pas de compétences particulières en évaluation. Au sein de la division de l'évaluation, il est plus particulièrement en charge de la coordination des évaluations décentralisées et de la mise à niveau du système d'information de la division de l'évaluation.

Emilie Aberlen, évaluatrice à l'AFD

Emilie Aberlen a été diplômée de Sciences po Paris en 1998. Elle est ensuite partie à Londres pour suivre une double maîtrise de droit des affaires français et anglais.

À la suite de cette formation, elle rejoint le cabinet d'évaluation Euréval où elle est chargée d'études. Elle y est seulement restée un an. En 2006, elle est engagée à Ernst and Young dans la division de l'évaluation des politiques publiques. Forte de cette expérience, elle part pour la Banque européenne d'investissement en 2012 où elle assure la fonction d'évaluatrice. Elle quitte la Banque seulement quelques mois plus tard. Après cette expérience, elle est embauchée à l'AFD au sein de la division de l'évaluation.

Corinne de Peretti, évaluatrice à l'AFD

Corinne de Peretti est économiste du développement, docteur en sciences économiques de l'Université Clermont I (CERDI). Elle a travaillé cinq ans au ministère des affaires étrangères sur les questions de gouvernance financière, d'intégration régionale et d'allègement de dette (C2D), puis a intégré l'AFD en 2008 pour occuper un poste de coordonnatrice géographique au département Afrique (Côte d'Ivoire, Nigeria, Liberia). Elle a rejoint en septembre 2012 la division Évaluation et capitalisation, où elle est notamment en charge de l'évaluation des aides budgétaires. Arrivée à la division d'évaluation, elle nous confie ne pas connaître grand-chose à l'évaluation. Elle est toutefois très volontaire pour se former et réaliser des évaluations « qui vont permettre un apprentissage et une amélioration des programmes d'aide de l'AFD.

Esmee Gaussen, évaluatrice au DFID

Esmee Gaussen est diplômée de l'université d'Oxford en sciences humaines. En 2002, elle suit une deuxième maîtrise d'anthropologie, d'environnement et de développement à *l'University College of London*. Elle fait ensuite un stage au Programme alimentaire mondial où elle est chargée du suivi des programmes d'aide. C'est durant ce stage qu'elle pratique l'évaluation pour la première fois et elle est tout de suite très intéressée par cette activité. En 2005, elle s'engage dans l'initiative *Student partnership worldwide* et part en mission en Tanzanie où elle est chargée du suivi et de l'évaluation des projets de coopération. En 2007, elle rentre au

Royaume-Uni et est embauchée au *National Audit Office* en tant que chargée de projet sur la coopération technique internationale. Elle se forme alors aux méthodes de mesure de l'efficience des programmes de coopération et devient spécialiste de la value for money. C'est à ce titre qu'elle est mutée au DFID. Fin 2011, ayant envie de renforcer son expérience à l'international, elle postule à Oxfam où elle est engagée pour réaliser une évaluation participative en Afrique du Sud sur l'éducation des filles. Cette mission achevée, elle retourne au DFID, où elle prend le poste de chargée de suivi-évaluation au sein de l'agence de Pretoria en Afrique du Sud. En 2013, elle souhaite rentrer au siège pour des raisons personnelles et rentre au sein du département de l'évaluation où elle est chargée des programmes sur le climat et l'environnement.

Penny Hawkins, directrice du département d'évaluation du DFID

Venue de Nouvelle Zélande, Penny Hawkins a été diplômée de la *Victoria University* de Wellington en psychologie au début des années 1990. Elle part ensuite pour le Royaume-Uni pour suivre un master sur les techniques de la communication à l'université de Keele.

Elle retourne en Nouvelle-Zélande et cherche un travail dans la fonction publique. Elle est embauchée au ministère du développement social en 1995 au sein de l'unité d'évaluation. Quand elle prend ce poste, elle n'a pas de compétences particulières en évaluation, mais elle s'engage rapidement dans les réseaux régionaux d'évaluation. Après être restée huit ans à ce poste, elle rentre au ministère des affaires étrangères et du commerce de la Nouvelle Zélande en tant que directrice de l'unité d'évaluation. Elle occupe ce poste pendant sept ans. Durant cette période, elle se distingue par son dynamisme dans la promotion d'une culture de l'évaluation dans la région Asie-Pacifique et devient présidente du réseau d'évaluateurs Australasia. Réputée comme étant une animatrice de réseau talentueuse, elle devient, en 2007, Vice Chair du réseau d'évaluateurs du CAD-OCDE.

En 2010, elle quitte la fonction publique et la Nouvelle Zélande et part aux États-Unis pour travailler en tant qu'évaluatrice à la Fondation Rockfeller. Elle s'engage également dans le comité de direction de *l'International Program for Development Evaluation* (IPDET). Elle contribue largement à diffuser une culture de l'évaluation au sein de la fondation Rockfeller et persuade ses dirigeants d'investir dans le renforcement des capacités en évaluation dans les pays du Sud. C'est ainsi que la fondation Rockfeller devient un des financeurs de l'initiative CLEAR. Fin 2013, Penny Hawkins est approchée pour succéder à Nick York en tant que directrice du département d'évaluation du DFID.

Joëlline Bénéfice, évaluatrice au CAD-OCDE

Joëlline Bénéfice est diplômée d'un Master de relations internationales de Sciences po Paris en 2006. Elle commence sa carrière en tant que responsable des relations extérieures au sein de l'ONG CCFD Terre solidaire au Sri Lanka. C'est durant cette expérience qu'elle est sensibilisée à l'évaluation des projets d'aide au développement. À la fin de son contrat, elle cherche donc

un poste dans un cabinet privé d'évaluation des politiques publiques et est embauchée au sein du cabinet luxembourgeois SEE. Elle y reste pendant plus de cinq ans. En 2012, elle rejoint la fonction publique française, où elle devient évaluatrice au ministère des affaires étrangères. Déçue par cette expérience, elle postule alors au CAD-OCDE où elle obtient un poste d'analyste en « résultats et évaluation du développement ». Elle est notamment chargée de coordonner le réseau d'évaluateurs du CAD-OCDE.

Bibliographie

RELATIONS INTERNATIONALES

Ouvrages

BADIE, Bertrand, *L'État importé. Essai sur l'occidentalisation de l'ordre politique,* Fayard, 1987, 334p

BATISTELLA, Dario, PETITVILLE, Frank, SMOUTS, Marie-Claude, *et al*. *Dictionnaire des relations internationales : approches, concepts, doctrines*, Dalloz, 2012, 572 p

COLONOMOS, Ariel (dir.), *Sociologie des réseaux transnationaux, Communautés, entreprises et individus : lien social et système international,* l'Harmattan, 2001, 300p

DEVIN, Guillaume, *Sociologie des relations internationales*, La Découverte, 2007, 128p

DEZALAY, Yves, GARTH, Bryant G, *Global Prescription. The Production, Exportation, and Importation of a New Legal Orthodoxy,* Ann Arbor, University of Michigan Press, 2002, 368p

HAAS, Peter, *Knowledge, Power and International Relations,* University of South Carolina Press, 1997, 404p

KECK, Margaret, SIKKINK, Kathryn, *Activist beyond borders: Advocacy Networks in International Politics,* Cornwell University Press, 1998, 240p

KLEIN, Asmara, LAPORTE, Camille, SAIGET, Marie, *Les bonnes pratiques des organisations internationales*, Presses de Sciences po, à paraître en mars 2015

KRASNER, Stephen D, *International Regimes*, Cornell University Press, 1985, ed. 1983, 372p

LUMSDAINE, David H, *Moral Vision in International Politics: The Foreign Aid Regime 1949-1989*, Princeton, Princeton University Press, 1993, 376p

MACDONALD, Donald S, *US-Korea relations from liberation to Self-Reliance: the twenty-year record*, Westview PR, 1992, 365p

Articles

BIGO, Didier, « Sécurité et immigration: vers une gouvernementalité par l'inquiétude », *Cultures et Conflits, Sécurité et Immigration*, n°31, 1998, pp. 13-38

HAAS, Peter, "Epistemic Communities and International Policy Coordination", *International Organization*, vol. 46, n°1, hiver 1992, pp. 1-35

KECK, Margaret, SIKKINK, Kathryn, “Transnational advocacy networks in international and regional politics”, *International Social Science Journal*, vol. 51, n°159, mars 1999, pp. 89-101

Littérature grise

Conseil européen, *Une Europe sûre dans un monde meilleur : Stratégie européenne de sécurité*, Bruxelles, 12 décembre 2003

HM Government, *Securing Britain in an Age of Uncertainty, The Strategic Defense and Security Review*, octobre 2010

Security Council of the United Nations, *Rapport du Secrétaire général sur les causes des conflits et la promotion d'une paix et d'un développement durables en Afrique*, 16 avril 1998

AIDE AU DEVELOPPEMENT

Ouvrages

ALMOND, Gabriel A, COLEMAN, James S, *The politics of the Developing Areas*, Princeton University Press, 1960, 608p

ALTANI-DUAULT, Vidal, *Anthropologie de l’aide humanitaire et du développement, des pratiques aux savoirs, des savoirs aux pratiques*, Armand Colin, 2009, 311p

BANERJEE, Abhijit Vinayak, *Making Aid Work*, Cambridge, Mass: the MIT Press, 2007, 170

CARMENT, David, PREST, Stewart, YIAGADEESEN, Samy, *Security, Development and the Fragile States: Bridging the gap between theory and policy*, Routledge Studies in Intervention and State building, 2010, 282p

CERNEA, Michael M, *la Dimension humaine dans les projets de développement*, Karthala, 1999, 586p

CHAMBERS, Robert, *Rural Development: Putting the Last First*, Longmans, 1983, 256p

CHATAIGNER, Jean-Marc, MAGRO Hervé, *États et sociétés fragiles : entre conflits, reconstruction et développement*, Éditions Karthala, 2007, 565p

DANN, Philip, *The Law of Development Cooperation, A Comparative Analysis of the World Bank, the EU and Germany*, Cambridge University Press, 2013, 604p

DUFFIELD, Mark, *Development, security and unending war: governing the world of people*, Cambridge Policy, 2007, 266p

DUFLO, Esther, *Expérience, science et lutte contre la pauvreté*, Fayard, 2009, 60p

EASTERLY, William, *The White Man's Burden: why the West's efforts to aid the rest have done so much ill and so little good*, The Penguin Press, 2006, 436p

FERGUSON, James, *The Anti-Politics Machine: Development, Depoliticization and Bureaucratic Power in Lesotho*, University of Minnesota Press, 1994, 336p

GABAS, Jean-Jacques, *Nord-Sud : l'impossible coopération*, Presses de Sciences po, 2002, 114p

GEROMINI, Vincent, *Savoirs et politiques de développement : questions en débat à l'aube du XXIème siècle*, Karthala-Gemdev, 2008, 253p

GROVES, Leslie, HINTON, Rachel, *Inclusive aid: changing power and relationships in international development*, Earthscan, 2004, 237p

GUENEAU, Marie-Christine, *Afrique, les petits projets de développement sont-ils efficaces ?* Editions L'Harmattan, Paris, 1986, 230p

HIRSCHMAN, Albert O, *Development Projects Observed*, The Brookings Institution, 1967, ed. 2002, 218p

LECOMTE, Bernard, NAUDET, Jean-David, *Survivre grâce à l'aide, réussir malgré l'aide*, Éditions de l'Aube, 2000, 211p

LEFTWICH, Adrian, *States of Development: on the Primacy of Politics in Development*, Polity Press, 2000, 240p

MAWDSLEY, Emma, *From Recipients to Donors: Emerging donors and the Changing Development Landscape*, Zed Books, 212, 288p

Mc NAMARA, Robert, *One Hundred Countries, Two Billions People: the dimensions of Development*, Praeger, 1973, 140p

MCEWAN, Cheryl, *Post colonialism and development*, Routledge, 2009, 376p

NEUMAYER, Eric, *The Pattern of Aid Giving: The Impact of Good Governance on Development Assistance*, Routledge, 2003, 128p

NUSSBAUM, Martha, SEN, Amartya, *The quality of life*, Oxford University Press, 1983, 472p

OLIVIER, Guillaume, (dir), *L'aide au développement, un outil à réinventer,* Éditions Charles Léopold Mayer, 2004, 177p

RIDDELL, Roger C, *Foreign Aid Reconsidered*, The John Hopkins University Press, 1987, 320p

RIDDELL, Roger, *Does Foreign Aid Really Work?* Oxford University Press, 2007, 536p

RIST, Gilbert, *Le développement, histoire d'une croyance occidentale*, Presses de Sciences po, 4ème ed., 2013, 512p

ROSTOW, William, *The Stages of Economic Growth: A Non-Communist Manifesto,* Cambridge University Press, 1960, third ed. 1991, 324p

SEN, Amartya, *On ethics and Economics*, Basil Blackwell, 1987, 192p

SEN, Amartya, *Development as freedom*, Oxford University Press, 1999, 384p

SEVERINO, Jean-Michel, DEBRAT, Jean-Michel, *L'aide au développement,* Le Cavalier Bleu, Paris, 2009, 127p

STIGLITZ, Joseph, *State versus Market : Have Asian currency crises affected the reform debate*, Published for the Bangladesh Economic Association 50p

STOKKE, Olav, *The UN and Development, from aid to cooperation*, Indiana University Press, 2009, 721 p

TANDON, Yash, *Ending Aid Dependence,* Fahamu, 2nd ed., 2008, 164p

WITHFIELD, Lindsay, *The Politics of Aid: African Strategies for Dealing with Donors*, Oxford University Press, 2009, 401p

Articles

ALESINA, Alberto, DOLLAR, David, "Who give Foreign Aid to Whom and Why ?", *Journal of Economic Growth,* vol. 5, 2000, pp. 33-63

ALESINA, Alberto, WEDER, Beatrice, "Do Corrupt Governments Receive Less Foreign Aid ?", *American Economic Review,* vol. 92, n°4, 2002, pp. 1126-1137

BAIRD, Ian G, SHOEMAKER, Bruce, "Unsettling Experiences: Internal Resettlement and International Aid Agencies in Laos", *Development and Change,* vol. 38, n°5, 2007, pp. 865-888

BOONE, Peter, "Politics and the effectiveness of foreign aid", *European Economic Review,* vol. 12, 1996, pp. 29-58

BURNSIDE, Craig, DOLLAR, David "Aid spurs growth – in a sound policy environment", *Finance and Development,* vol. 34, n°4, 1997, pp. 4-7

BALLEIX, Corinne, « La politique française de coopération au développement: Cinquante ans d'histoire au miroir de l'Europe », *Afrique Contemporaine,* vol. 4, n°236, 2010, pp. 95-107

CALVES, Anne-Emmanuèle, « Empowerment : généalogie d'un concept clé du discours contemporain sur le développement », *Revue Tiers monde,* vol.4, n°200, 2009, pp. 735-749

CHATAIGNER, Jean-Marc, « aide publique au développement et réformes des systèmes de sécurité : l'improbable rencontre du Dr Jeckyll et de Mr Hyde », *Afrique contemporaine,* n°209, janvier 2004, pp. 39-49

CHAUVET, Linda, COGNEAU, Denis, NAUDET, Jean-David, « Sélectivité et égalité des chances dans l'allocation de l'aide internationale. Une analyse de la dernière décennie », Économie et prévision, n°186, 2008, pp. 23-38

CHENERY, Hollis B, STROUT Alan, "Foreign Assistance and Economic Development", *American Economic Review*, n°56, 1966, pp. 680-733

CHUN, Hong Mi, MUNYI, Elijah. N, LEE, Heejin, "South Korea as an Emerging donor: challenges and change on its entering OECD-DAC", *Journal of International Development,* vol. 22, 2010, pp. 788-802

COLLIER, Paul, "African Growth: Why a Big-Push?" *Journal of African Economies,* vol. 15, 2006, pp. 188-211

DENIZER, Cevdet, KAUFMANN, Daniel, KRAAY, Aart, "Good countries or good projects? Macro and micro correlates of World Bank project performance", *Journal of Development Economics,* vol. 105, novembre 2013, pp. 288-302

EASTERLY, William, "Can Foreign Aid Buy Growth", *The Journal of Economic Perspectives,* vol. 17, n°3, 2003, pp. 23-48

EYBEN, Rosalind, "Struggles in Paris: the DAC and the Purpose of Development Aid", *European Journal of Development Studies,* n°25, 2013, pp. 78-91

EYBEN, Rosalind, "Emerging and Submerging Powers: Imagined Geographies in the New Development Partnership at the Busan Fourth High Level Forum", *The Journal of Development Studies,* vol. 49, n°4, 2013, pp. 457-46

GABAS, Jean-Jacques, « La mesure du développement : comment science et politique se conjuguent », *Revue Tiers Monde,* n°213, janvier 2013, pp. 1-11

GABAS, Jean-Jacques, HUGUON, Philippe, « Les biens publics mondiaux et la coopération internationale », *L'Économie politique,* vol. 4, n°12, 2011, pp. 19-31

GABAS, Jean-Jacques, LAPORTE, Camille, « La pauvreté dans l'agenda des Nations unies », *CERISCOPE Frontières,* 2011

GIOVALUCCHI, François, OLIVIER DE SARDAN, Jean Pierre, « planification, gestion et politique dans l'aide au développement : le cadre logique, outil et miroir des développeurs », *Revue Tiers Monde*, vol. 2, n°198, 2009, pp. 383-406

GOUNDER, Rukami, "What Motivates Foreign Aid : A Case Study of Australia Aid to Indonesia", *Journal of Developing Areas,* vol. 33, n°3, 1999, pp. 379-394

GRIFFIN, Karl, ENOS, John, "Foreign Assistance: Objectives and Consequences", *Economic Development and Cultural Change,* vol. 18, n°3, 1970, pp. 313-327

HATTORI, Tomohisa, "The Moral Politics of Foreign Aid", *Review of International Studies,* vol. 29, n°2, avril 2003, pp. 229-247

HUANG, Meibo, REN, Peiqiang, « L'aide étrangère de la Chine dans l'architecture de l'aide internationale », *Revue internationale de politique de développement*, vol. 3, 2012, pp. 1-18

KIM, Eun Mee, KIM, Pil Ho, KIM Jin Kyung, "From development to development cooperation: foreign aid, country ownership and the developmental state in South Korea", *The Pacific Review,* vol. 26, n°3, février 2013, pp. 313-336

KIM, Sung-Mi, MAWDSLEY, Emma, SAVAGE, Laura, "A post aid world? Paradigm shift in foreign aid and development cooperation at the 2011 Busan High Level Forum", *The Geographical Journal,* vol. 180, n°1, mars 2014, pp. 27-38

LENSINK, Robert, MORISSEY, Olivier, "Aid Instability as a Measure of Uncertainty and the Positive Impact of Aid on Growth", Journal of Development Studies, vol. 36, n°3, 200, pp. 31-49

MARX, Axel, SOARES, Jadir, "South Korea's Transition from Recipient to DAC Donor: Assessing Korea's Development Cooperation Policy", *International Development Policy,* 2013, pp. 107-142

MANNING, Richard, "Will Emerging Donors' Change the Face of International Co-operation?" *Development Policy Review,* vol. 24, n°4, juillet 2006, pp. 371-385

MOSLEY, Paul, "The Political Economy of Foreign Aid: A Model of the Market for a Public Good", *Economic Development and Cultural Change,* vol. 33, n°2, janvier 1985, pp. 373-93

PARIZET, Raphaëlle, « Mesurer le développement pour gouverner les peuples autochtones », *Revue Tiers Monde,* n°213, janvier 2013, pp. 143-160

POUILLIEUTE, Antoine, « la lutte contre la pauvreté, nouveau paradigme de l'Agence française de développement », *Politique africaine,* n°81, 2001, pp. 171-175

ROSENSTEIN-RODAN, Paul, "The International Development of Economically Backward Areas", *International Affairs*, vol. 20, n°2, avril 1944, pp. 157-165.

SERRE, Julien, « Aide au développement et lutte contre le terrorisme », *Politique étrangère,* hiver 2012, pp. 891-904

SEVERINO, Jean-Michel, « Vers des objectifs mondiaux pour la globalisation », *CERISCOPE Pauvreté,* 2012

SMITH, Kimberley, ZIMMERMANN, Felix, "More Actors, More Money, More Ideas for International Development Cooperation", *Journal of International Development,* vol. 23, n°5, juillet 2011, pp. 722-738

SVENSSON, Jakob, "Why Conditional Aid Does not Work and What Can Be Done About It ?", *Journal of Development Economics,* vol. 70, n°2, 2000, pp. 381-402

VAN WAEYENBERGE, Elisa, "Selectivity at Work : Country Policy and Institutional Assessments at the World Bank", *European Journal Development Research*, 2009, pp.792-810

WOODS, Ngaire, “Whose aid? Whose influence? China, emerging donors and the silent revolution in development influence” *International Affairs*, vol. 85, n°6, 2008, pp. 1205-1221

Littérature grise

Ausaid, *Mutual Accountability: « orphan » principle of the Paris Declaration*, ODE Issues Note, juin 2008

BRÄUTIGAM, Deborah, *China, Africa and the International Aid Architecture*, working paper, Banque africaine de développement, mars 2010

Busan High Level Forum on aid Effectiveness, Proceedings, décembre 2011

CHAMBERS, Robert, *Farmers First: A Practical Paradigm for Third World Agriculture*, Institute of Development Studies, 1988

China’s Information Office of the State Council, *White Paper on China’s Foreign Aid,* 21 Avril 2011

COGNEAU, David, NAUDET, Jean-David, *Qui mérite l’aide ? Égalité des chances versus sélectivité*, AFD, 2004

COLLIER, Paul, DOLLAR, David, Aid Allocation and Poverty Reduction, World Bank, Development Research Group, avril 1999

Conseil d’État, *Décret n° 2009-618 du 5 juin 2009 relatif à l'Agence française de développement*, article 2, juin 2009

DAC-OECD, *Investir dans le Tiers monde*, 1972

DAC-OECD, *Rapport annuel sur la coopération pour le développement*, 1965

DAC-OECD, *Revue par les pairs : Suède*, 1985

DAC-OECD, *Glossaire du CAD de l’OCDE sur les principaux termes relatifs à l’évaluation*, 1991

DAC-OCDE, *Shaping the 21st Century: The Contribution of Development Cooperation*, 1996

DAC-OECD, *Development Cooperation Review: United Kingdom,, 1998*

DAC-OECD, *Development Cooperation Review: France : L'élimination de la pauvreté dans le monde : Faire de la mondialisation un succès pour les pauvres*, 2000

DAC-OECD, *Harmonizing donor practices for effective aid delivery*, 2003

DAC-OECD, *Inscrire la coopération pour le développement dans une optique de prévention du terrorisme*, 2003

DAC-OECD, *Déclaration de Paris sur l’aide au développement,* 2005

DAC-OECD, *Historique du Comité d'aide au développement de l'OCDE,* 2006

DAC-OECD, *Programme d'action d'Accra*, 2008

DAC-OECD, *Mutual Accountability : an emerging good practice*, 2009

DAC-OECD, *Conflits et fragilité. Le Financement des États en transition : vers une meilleure réponse*, 2010

DAC-OECD, *Peer Review: United Kingdom*, 2010

DAC-OECD, *Ownership and Mutual Accountability, 10 top tips on mutual accountability,* 2010

DAC-OECD, *Partenariat de Busan pour une coopération efficace au service du développement*, 2011

DAC-OECD, *Peer review : France*, 2013

DAC-OECD, *Identification and Monitoring of Potentially Under-aided Countries*, 2013

DAVIES, Penny, *Aid Effectiveness and non-DAC providers of Development Assistance,* Informal Working Group on non-DAC Providers of Development Assistance Background document to Round Table 9 of the Third High Level Forum on Aid Effectiveness, août 2008

DFID, *Fighting poverty to build a safer world, a strategy for security and development*, mars 2005

DFID, *Eliminating World Poverty Building our Common Future,* White Paper, juillet 2009

DFID, UKAID, *Bilateral Aid Review: Technical report*, mars 2011

DFID, *DFID's Approach to Value for Money (VfM),* juillet 2011

ECA-OECD, *Development Effectiveness in Africa, Promise and performance: Applying mutual accountability,* octobre 2005

ENLAI, Zhou, *Eight Principles for China's Aid to Foreign Countries,* 1964

FUES, Thomas, KLINGEBIEL, Stephan, *Unrequited love: What is the legacy of the first Global Partnership summit,* The Current Column, DIE-GDI, 2014

GELB, Alan, NGO, Brian, YE, Xiao, *Implementing performance-based aid in Africa: The country policy and institutional assessment,* World Bank Africa Region, Working Paper 77, 2004

The Global Fund, *Overview of the Allocation Methodology (2014-2016), The Global Fund's new funding model,* mars 2014

Government of the Lao People's Democratic Republic and the Partners in Development, *Vientiane Declaration on Aid Effectiveness,* 2007

GRAVES, Sue, *Mutual accountability at the country level, Vietnam country case study,* ODI-OECD, Joint Venture on Managing for Development Results, 2009

HADJIMICHAEL, Michael T, *et al*., *Sub-Saharan Africa: Growth, savings, and investment, 1986-93,* Occasional Paper 118, International Monetary Fund, 1995

HAMMARSKJÖLD, Dag, *Rapport Dag Hammarskjöld : que faire : un autre développement,* préparé à l'occasion de la septième assemblée extraordinaire des Nations unies, 1975

Haut Conseil de la Coopération Internationale, *Les priorités de la coopération pour l'Afrique subsaharienne et le nouveau partenariat pour le développement de l'Afrique (NEPAD),* rapport, Avril 2002

Independent Commission for Aid Impact, *How DFID learns,* report 34, 2014

International Aid Transparency Initiative Accra Statement, Ghana, 4 septembre 2008

International Development (Reporting and Transparency) Act, 2006

KOICA-UNRISD, *Making International Development Cooperation Effective,* 2011

Ministères des affaires étrangères, *stratégie de coopération au développement,* Paris, 2007

Ministère des affaires étrangères, *Coopération au développement : une vision française,* document cadre, Direction générale de la mondialisation, 2011

Ministère des affaires étrangère, *Stratégie genre et développement 2013-2017,* Direction générale de la mondialisation, 2013

MOSLEY, Paul, Aid as obstacle: Twenty questions about our foreign aid and the hungry, Institute for Food and Development Policy, 1980

NAUDET, Jean-David, *Les OMD et l'aide de cinquième génération : Analyse de l'évolution des fondements éthiques de l'aide au développement, Agence française de développement*, document de travail, 2005

NORAD, *Democracy Support through the United Nations,* Evaluation Department, octobre 2010

People's Republic of China, *China's Peaceful Developmen*t. Beijing, China: Information Office of the State Council, 2011

RAJAN, Raghuram G, SUBRAMANIAN, Arvind, *Aid and Growth: What Does the Cross-Country Evidence Really Show,* IMF Working Paper n°127, 2005

SACHS, Jeffrey, *Investir dans le développement : plan pratique pour réaliser les Objectifs du Millénaire pour le développement,* Millenium Projet, 2005

SADEV, *Mutual accountability in practice, the case of Mozambique*, juin 2012

Secretary of State for International Development, *White Paper on International Development: Reducing poverty a challenge for the 21st century*, novembre 1997

Secretary of State for International Development, *White Paper on International Development: Eliminating World Poverty: Making globalization works for the poor*, 2000

SMITH, Kimberley, YAMASHIRO, Tallta, ZIMMERMANN, Felix, *Beyond DAC, the Welcome role of other providers of development cooperation*, OECD Development Cooperation Directorate, mai 2012

TALAAT, Abdel-Malek, *The Global Partnership for Effective Development Cooperation: What next for a global architecture*, The Current Column, DIE-GDI, 2002

TSIKATA, Tsidi, *Aid effectiveness: A survey of the recent empirical literature*, International Monetary Fund paper on policy analysis and assessment, 1998

UNDP, *Rapport mondial sur le développement humain : Le développement humain au service de l'éradication de la pauvreté*, 1997

United Nations, *Monterrey Consensus on Financing for Development*, 2003

VERDELHAN-CAYRE, Geneviève, *L'AFD et la prise en compte du genre*, Note opérationnelle, AFD, janvier 2001

WOLFENSOHN, James, *People and Development. Annual Meeting Address*, World Bank, octobre 1996

World Bank, *Annual Report of Project Performance Results*, Report n°5859, OED, 1985

World Bank, *Transcript of the Interview with Mervin Weiner*, Oral History Program, The World Bank / IFC Archives, 1986.

World Bank, « Poverty », in *World Development Report 1990*, New York, 1990

World Bank, *World Development Report 1997 : The State in a Changing World*, 1997.

World Bank, *Assessing Aid: What works, what doesn't and Why?* New-York: Oxford University Press, 1998

World Bank, *Transcript of the interview with Yves Rovani*, Oral History Program, The World Bank / IFC Archives 2003

World Bank, *Transcript of the interview with Graham Donaldson*, Oral History Program, The World Bank / IFC Archives, 2012

Papiers de recherche

BEN-ARTZI, Ruth, *Ios and Peer Pressure: an Examination of the Development Assistance Committee (DAC),* Prepared for 8th Pan European Conference on International Relations, Warsaw, 21 septembre 2013

CANAVIRE, Gustavo, et al., *Assessing the Allocation of Aid: Development Concerns and the Self-Interest Donors*, working paper n°1253, Kiel Institute for World Economics, 2005

CASSEN, Robert, *Does Aid Work? Report to an international task force,* Oxford Clarendon press, 1986

CASTELLANET, Christian, « Cycle des projets, cadre logique et efficacité des interventions de développement », *Transverses*, n°13, GRET, 2003

CHAUVET, Linda, GUILLAUMONT, Pierre, "Aid and Growth Revisited: Policy, Economic Vulnerability, and Political Instability", in TUNGODDEN, Bertil, STERN, Nicholas, KOLSTAD, Ivar (dir. pub.), *Towards Pro-Poor Policies – Aid, Institutions and Globalization*, World Bank/Oxford University Press, 2004

CLEMENS, Michael, RADELET, Steven, BHAVNANI, Rikhil, *Counting Chickens when they hatch: the short-term effect of aid on growth,* Working Papers 44, Center for Global Development, 2004

DUFLO, Esther, in MANERY, Florian, « L'économie du développement à l'épreuve du terrain, Entretien avec Esther Duflo », *la vie des idées*, 5 mai 2009

DURBARRY, Ramesh, GEMMELL, Norman, GREENAWAY, David, *New Evidence on the Impact of Foreign Aid on Economic Growth,* Centre for Research in Economic Development and International Trade, University of Nottingham, 1998

EASTERLY, William, *The cartel of good intentions: the problem of bureaucracy in foreign aid,* Center for Global Development, octobre 2002

EIFERT, Benn, GELB, Alan, *Improving the dynamics of aid: towards more predictable budget support,* World Bank Policy Research Working Paper, n°3732, 2005

GUILLAUMONT, Patrick, *Adapting Aid Allocation Criteria to Development Goals,* An essay for the 2008 Development Cooperation Forum United Nations Economic and Social Council, mai 2008

GRIMM et al, *European Development Cooperation to 2020: Challenge by New Actors in International Development,* Working Paper, EDC 2020, 2009

HAFTERDORN, Esther, *État des lieux du secteur de l'hydroélectricité au Laos et identification des opportunités de financement pour l'AFD,* AFD, Bureau de Vientiane, RDP Lao, novembre 2013

HANSEN, Henrik, TARP, Finn, *Aid and growth regressions,* DERG, University of Copenhagen, 1999

HOEFFLER, Anke, OUTRAM, Verity, *Need, Merit or Self Interest – What Determines the Allocation of Aid*, Oxford University, Department of Economics, juillet 2008

HULME, David, *The Making of the Millennium Development Goals: Human Development Meets Result-Based Management in an Imperfect World*, University of Manchester, Institute for Development Policy and Management, 2007

ISOPI, Alessia, MAVROTAS, George, *Aid Allocation and Aid Effectiveness, an Empirical Analysis*, United Nations University, UNU-WIDER, janvier 2006

LEE, Seong Hoon, HONG Moon Suk, KANG Hanee, *The Outcomes and Future Challenges of the Busan HLF-4: CSO's Perspectives*, Issue Brief 2, KOFID, REDI, février 2012

MANNING, Richard, *The DAC as a central actor in development policy issues: experiences over the past four years,* Discussion Paper, Deutsches Institut für Entwicklungspolitik, 2008

PARK, Kang Ho, "Korea's Role in Global Development", *Brookings,* février 2010

PETIT, Bernard, *La banalisation technocratique de l'aide*, document de travail, Série « Politiques de développement », FERDI, mars 2011

ROODMAN, David, *Macro Aid Effectiveness Research: A Guide for the Perplexed,* Center for Global Development, décembre 2007

SCHULZ, Nils-Sjard, *From Accra to 2011: perspective for the global governance of aid,* FUES, 2011

STEER, Liesbet, WATHNE, Cecilie, *Mutual Accountability at country level: Emerging good practice,* ODI, Background Note, avril 2009

STEIN, Danielle, VALTERS, Craig, *Understanding Theory of Change in International Development,* the Justice and Security Research Programme, London School of Economics, août 2012

VIELAJUS, Martin, *et al.*, *Le défi de la « redevabilité » des agences de développement dans le pays et dans leurs propres pays et face à leurs pair, une étude comparative »,* AFD, IRG, GRET, ODI, 2010

WILD, Leni, ELHAWARY, Samir, *The UK's approach to linking development and security: assessing policy and practice,* ODI, Working paper 347, mai 2012

Thèses et mémoires

De ALMEIDA CRAVO, Teresa Paula, *What's in a label, the aid community's representation of success and failure in Mozambique and Guinea Bissau,* Department of Politics and International Studies, University of Cambridge, Supervisor Devon Curtis, janvier 2012, 271p

BEDECARRATS, Florent, *La microfinance entre utilité sociale et performances financières : le rôle des normes dans la gouvernance d'un secteur mondialisé,* Thèse dirigée par Madame le

professeur Josepha Laroche, Université Paris I Panthéon Sorbonne, École doctorale de science politique, 22 juin 2012, 418p

BRANT, Philippa, *No Strings Attached: Chinese Foreign Aid and its Implications for the International Aid Regime,* Submitted in total fulfilment of the requirements of the degree of Doctor of Philosophy, School of Social and Political Sciences, Faculty of Arts, The University of Melbourne, mai 2012, 273p

CORTINOVIS, Mathilde, *Engagements et résistances au genre dans le développement : le cas de l'Agence française de Développement, 1997-2013*, sous la direction de la Professeure Delphine Dulong, Master Coopération internationale, Action humanitaire et Politiques de Développement, Université Paris 1, décembre 2013, 78p

DESCHAMBAULT, Joëlle, *Sécurité et développement dans la politique de coopération internationale: une approche biopolitique. Le cas du Canada en Haïti (1994-2008)*, mémoire de Maîtrise en Sciences-Politiques, Université du Québec à Montréal, septembre 2009, 144p

GOYETTE-COTE, Gabriel, *Le « nouveau régime de l'aide » à l'épreuve des faits : du Consensus de Washington à la Déclaration de Paris,* Mémoire de Science politique, Université du Québec à Montréal, mai 2010, 118p

YAMAMOTO, Yumiko, *L'aide international, au-delà des questions techniques: les politiques d'aide au développement à destination de l'Afrique de l'Ouest de trois pays asiatiques: Chine, Inde, Japon (1997-2010*), Thèse de doctorat en Science politique, dirigée par Madame Catherine Withol de Wenden et Monsieur Pierre Lascoumes, Institut d'Études Politiques de Paris, 2012, 369p

Résolutions des Nations unies

Assemblée générale des Nations unies, *résolution 3281*, 12 décembre 1974

Allocutions

Hilary Cliton, *Allocution lors de la session pleinière d'ouverture du Haut Forum de Busan, 29 novembre 2011*

Graham Donaldson, *Allocution lors du séminaire organisé par la Banque mondiale pour les 30 ans de son département d'évaluation*, archives de la Banque mondiale

Hadji Gueye, Docteur en économie, Coordonnateur du Centre CLEAR de l'Afrique francophone, allocution lors d'un séminaire de la société française d'évaluation intitulée « quelles perspectives pour l'évaluation de l'aide au développement » organisée le 13 février 2013 à Paris

Patrick G Grasso, *Allocution lors du séminaire de la Banque mondiale célébrant les 30 ans de son département d'évaluation*, archives de la Banque mondiale.

Prem Garg , Joanne Salop, Jan Piercy, *Allocution lors du séminaire de la Banque mondiale célébrant les 30 ans de son département d'évaluation*, archives de la Banque mondiale

Marco Farani, *Allocution au Forum de Busan, section thématique sur la coopération Sud-Sud et triangulaire*, 29 novembre 2011

Susan Stout, *Allocution lors du séminaire de la Banque mondiale célébrant les 30 ans de son département d'évaluation,* archive de la Banque mondiale

Harry S Truman, *Speech to the Congress*, Washington, March 12th 1947

Mervin Weiner, *Allocution lors du séminaire des 30 ans du département d'évaluation de la Banque mondiale*, archive de la Banque mondiale.

EVALUATION DE L'AIDE AU DEVELOPPEMENT

Ouvrages

BERLAGE, Lodewijk, STOKKE, Olav, *Evaluating Development Assistance: Approaches and Methods*, EADI Book Series 14, Frank Cass,1992, 219p

BORUCH, Robert, *Place randomized Trials: Experimental Test of Public Policy,* The annals of the American Academy of Political and Social Science, Philadelphia, 2005, 286p

BOYLE, Richard, LEMAIRE, Donald, *Building Effective Evaluation Capacity: Lessons from Practice*, Comparative Policy Evaluation, Transaction Publisher, février 1999, 202p

CRACKNELL, Basil. E, *Evaluating Development Aid: Issues, Problems and Solutions*, Sage Publication, 2000, 386p

DALE, Reidar, *Evaluating Development Programmes and Projects,* Sage Publication, 1998, 220p

HORTON, Douglas, *et al*, *Evaluation Capacity Development, Experience from Research and Development Organizations around the World*, ISNAR, IDRC, CTA, 2003, 188p

JACKSON, Edward T, KASSAM, Yusuf, *Knowledge Shared: Participatory Evaluation in Development Cooperation*, IDRC, Kumarian Press, 1998, 252p

LAPORTE, Camille, *Évaluation et efficacité de l'aide au développement en Corée du Nord*, L'Harmattan, 2011, 220p

REBIEN, Claus C, *Evaluating Development Assistance in Theory and in Practice,* Avedbury, England, 1996, 208p

RIST, Ray, BOILY, Marie-Hélène, MARTIN, Frederick R, *Development Evaluation in Times of Turbulence, Dealing with Crisis that Endanger our future,* The World Bank, 243p

Articles

BIRDSALL, Nancy, “Aid, Weak Institutions and the Missing Middle in Africa”, *Development Policy Review,* vol. 25, n°5, 2007, pp. 575-598

BOURGUIGNON, François, SUNDBERG, Mark, “Aid Can Work”, *Finance and Development,* vol. 44, n°1, mars 2007, pp. 1-12

BROWNING, Richard, “Evaluation in the ODA. A view from the inside”, *Public Administration and development,* vol. 4, 1984, pp. 133-139

CHERVEL, Marc, « La méthode des effets : présentations récentes, nouvelles citriques », *Tiers-Monde,* tome 22, n°87, 1981, pp. 603-612

CHERVEL, Marc, « La méthode des effets trente ans après », *Tiers-Monde,* tome 33, n°132, 1992, pp. 873-891

CLEMENTS, Paul, “Monitoring and Evaluation for Cost-Effectiveness in Development Management”, *Journal of Multidisciplinary Evaluation,* 2005, pp. 11-38

CONLIN, Sean, STIRRAT, Roderick, “Current challenges in development evaluation”, *Evaluation,* vol. 14, n°193, 2008, pp. 193-208

CRACKNELL, “Learning lessons from experience: the role of evaluation in the administration of the UK aid programme”, *Public administration and development,* vol. 4, 1984, pp. 15-20

CRACKNELL, Basil E, “Evaluating development assistance: a review of the literature”, *Public Administration and Development,* vol. 8, 1988, pp. 75-83

EASTERLY, William, PFUTZE, Tobias, “Where Does the Money Go? Best and Worst Practices in Foreign Aid”, *The Journal of Economic Perspectives,* vol. 22, n°2, printemps 2008, pp. 29-52

GRASSO, Patrick G, “What Makes an Evaluation Useful? Reflections from Experience in Large Organizations”, *American Journal of Evaluation,* vol. 24, 2003, n°4, pp. 507-514

KHAN, Adil, “Evaluation Capacity Building: An Overview of Current Status, Issues and Options”, *Evaluation,* vol. 4, n°3, 1998, pp. 310-328

KOSHELEVA, Natalia, SEGONE, Marco, “Evalpartners: Facilitating the Development of a New Model of Voluntary Organization for Professional Evaluation to Support the Development of National Evaluation Capacities”, *American Journal of Evaluation,* 30 juillet 2013, pp. 1-6

LABROUSSE, Agnès, « Nouvelle économie du développement et essais cliniques randomisés : une mise en perspective d’un outil de preuve et de gouvernement », *Revue de la régulation,* n°7, 2010

MICHAELOWA, Katharina, BORRMANN, Axel, “Evaluation Bias and Incentive Structures in Bi-and Multilateral agencies”, *Review of Development economics,* vol. 10, n°6, 2006, pp. 313-329

NAUDET, Jean-David, DEALRUE, Jocelyne, BERNARD, Tanguy « Évaluations d'impact, un outil de redevabilité ? Les leçons tirées de l'expérience de l'AFD », *Revue d'économie du développement*, n°4, vol. 16, 2012, pp. 27-48

SASAKY, Ryoh, "A Review of the History and the Current Practices of Aid Evaluation", *Journal of Multidisciplinary Evaluation*, n°5, 2006, pp. 55-83.

SCHAUMBÜRG, Miller, "Use of aid evaluation from an organizational perspective", *Evaluation*, vol. 11, n°2, pp. 207-222

Littérature grise

AFD, *Les évaluations sont-elles utiles ? Revue de littérature sur connaissances et décisions*, janvier 2009

AFD, *Bilan de l'assistance technique à la Fédération des paysans du Fouta Djallon (FPFD) en Guinée : 15 ans d'accompagnement*, Division évaluation et capitalisation, Edition Ex Post, n°32, août 2010

AFD, *Politique d'évaluation de l'AFD,* octobre 2013

Australian Government, Ausaid, *Australian Multilateral Assessment*, mars 2012

Ausaid, *2013 Quality Review of Aid Program Performance Reports*, Office of Development Effectiveness, janvier 2014

Australian Government, Ausaid, *Australian Multilateral Assessment*, mars 2012

BALOGUN, Paul, *Evaluation of MOPAN*, Swedish Institute for Public Administration, août 2013

BORMANN, Axel, STOKHMANN, Reinhard, *Evaluation in German Development Cooperation, A System Review*, Evaluation report 044, Federal Ministry for Economic Development Cooperation, 2010

CLAPP-WINCEK, Cynthia, BLUE, Richard, *Evaluation of USAID's Recent Evaluation Experience*, USAID Working Paper n°320, juin 2011

DAC-OECD, *Lessons of Experience*, Working group on Aid Evaluation, 1984

DAC-OECD(a), *Rapport de bilan sur l'évaluation dans les agences d'aide*, 1986

DAC-OECD(b), *Méthodes et procédures d'évaluation de l'aide : un recueil des pratiques des donateurs*, 1986

DAC-OECD(c), « Glossaire des termes utilisés dans l'évaluation », dans *Méthodes et procédures dans l'évaluation de l'aide,* 1986

DAC-OECD, Working Group on Evaluation, *Review of DAC principles for development assistance*, 1998

DAC-OECD Working Group on Aid Evaluation, *Result based management in development aid agencies: a review of experiences*, 1999.

DAC-OECD (b), *Fact-finding survey of evaluation capacity development (ECD) in partner countries*, octobre 2006

DAC-OECD, *Promoting Pro-Poor Growth, Practical Guide to Ex Ante Poverty Impact Assessment*, 2007

DAC-OECD, *Better Aid: Evaluation in Development Agencies*, 2010

DAC-OECD Network on Development Evaluation, *Assessing the Development Effectiveness of Multilateral Organizations: Guidance on the Methodological Approach*, juin 2012

DAC-OECD(a), *The DAC Network on Development Evaluation – 30 years of strengthening learning in development*, juin 2013

DAC-OECD, Evalnet, *Development Effectiveness Review of the World Health Organization, 2007-2010*, mars 2013

DAC-OECD, Evalnet(b), *Study on Collaborative Partner-Donor Evaluation Work*, Concept Note, document de travail, 26 juin 2013

DANIDA, *Joint Evaluation of the Road sub-sector Programme 1996-2000 in Ghana*, 2000

DANIDA, *Joint Government-Donor Evaluation of Basic and Primary Education in Nepal*, 2004

DANIDA, Joint Evaluation of Uganda's Plan for the Modernization of Agriculture, 2005

DANIDA, *Evaluation of Danish Assistance to Uganda 1987-2005*, 2006

DANIDA, *Capacity Development Outcome Evaluation of Danish Supported Organizations in Bhutan*, 2006

DANIDA, *Joint External Evaluation, the Health Sector in Tanzania 1999-2006*, 2007

DANIDA, *Joint Evaluation of The Danish Country Programme in Benin 1990-2006*, 2008

DANIDA, *Evaluation of Development Cooperation between Mozambique and Denmark 1992-2006*, 2008

DANIDA, *Joint Evaluation of the Ghana-Denmark Development Cooperation* 1990-2006, 2008

DFID, *Multilateral Aid Review, Ensuring maximum value for money for UK aid through multilateral organizations*, mars 2011

DFID, *International Development Evaluation policy*, mai 2013

DFID, *Multilateral Review Update, Driving reform to achieve multilateral effectiveness,* décembre 2013

Evaluation Capacity Development Group, *An International Workshop Agreement on Evaluation Capacity Development*, Genève, octobre 2011

FAO, *Participatory Assessment, Monitoring and Evaluation: A Field Manual. Rome, Community Forestry Unit*, FAO, 1989

FORESTI, Maria, et al., *A comparative Study of Evaluation Policies and Practices in Development Agencies,* AFD/ODI, décembre 2007

FORSS, Kim, *Participatory Evaluation: Questions and Issues*, a Report Commissioned by the Central Evaluation Office, Occasional Paper n°1, UNDP, octobre 1989

GUERIN, Isabelle, *et al.*, *Analyse des déterminants de la demande de services financiers dans le Maroc rural,* AFD, Département de la recherche et de l'évaluation, 2012

HAYES, Samuel, *Evaluating Development Projects*, UNESCO, 1959

ILO, *Response by the ILO to the UK-DFID Multilateral Aid Review Report,* 2011

JONES, Harry, MENDIZABAL, Enrique, *Strengthening learning from research and evaluation: going with the grain,* ODI, septembre 2010

KNACK, Stephen, ROGERS, Halsey, EUBANK, Nicholas, *Aid Quality and Donors Rankings,* Policy Research Paper, The World Bank, Development Research Group, Human Development and Public Service Team, Mai 2010

KUMAR, Krishna, *Conducting key informant interview in developing countries,* USAID, 1989

LEVINE, David, POLIMENI, Rachel, RAMAGE, Ian, *Insuring Health or Insuring Wealth? An Experimental evaluation of Health Insurance in Cambodia*, AFD, Département de la recherche et de l'évaluation, 2012

MELANSON, James, Development Evaluation Division, Foreign Affairs, Trade and Development Canada, *Comparative Study of MOPAN and Evalnet Approaches to Assessing Multilateral Organizations' Development Effectiveness,* février 2014

Ministry of Foreign Affairs of Denmark, Danida, *Experience with conducting evaluation jointly with partner countries,* Danida's Evaluation Department, juin 2009

PITMAN, George K, FEINSTEIN, Osvaldo N, INGRAM, Gregory K, *Evaluating Development Effectiveness,* World Bank Series on Evaluation and Development, n°7, 1997

Proceedings of the 1985 International Conference on Rapid Rural Appraisal, Khon Kaen University, 1987

ROBB, Caroline M, *Can the Poor Influence Policy? Participatory Poverty Assessment in Developing World,* Directions in Development, World Bank, 1999

RUGH, Jim, SEGONE, Marco, *Voluntary Organizations for Professional Evaluation (VOPEs), Learning from Africa, Americas, Asia, Australasia, Europe and Middle East, UNICEF*, Evalpartners, IOCE, 2013

Société française d'évaluation, *Varia : L'évaluation pour l'efficacité de l'aide au développement,* les cahiers de la SFE, janvier 2013

SIDA, *Are SIDA's evaluations good enough? An assessment of 34 evaluations report,* Sida Studies in Evaluation, janvier 2008

UNDP, *Who are Question-makers? A Participatory Evaluation Handbook,* Office of Evaluation and Strategic Planning, New York, 1997

UNPD, *Evaluation capacity development in Asia, Selecting Proceedings from the International Conference Beijing,* octobre 1999

UNEDAP, *Regional Thematic Study of national monitoring and evaluation systems in Asia and the Pacific,* Thaïlande, 2012.

UNEG, *National Evaluation Capacity Development: Practical tips on how to strengthen National Evaluation Systems,* United Nations, 2012

UNEP, *DFID Multilateral Aid Review UNEP Right to Reply*, 2011

UNESCO *Response to the UK Multilateral Aid Review*, 2011

UNICEF, *How to design and manage equity-focused evaluations*, 2011

UNICEF, IPEN, *New Trends in Development Evaluation,* Evaluation working papers, issue 5, 2006

UNIFEM, *UNIFEM works on gender-responsive budgeting,* Evaluation Unit, 2009

USAID, *Evaluation: Learning from Experience,* USAID Evaluation Policy, janvier 2011

WOOD, Brian, *et a*l., *L'évaluation de la Déclaration de Paris, Phase 2 rapport final,* DAC-OECD, mai 2011

World Bank, *World Bank Participation Sourcebook*, OED, 1996

World Bank, *Evaluation Capacity Development, Lessons from National Experience,* ECD Working Papers serie n°1, OED, mai 1998

World Bank, *Evaluation and Poverty Reduction,* proceedings of a World Bank Conference, OED, 1999

World Bank, *Building Country Capacity for Monitoring and Evaluation in the Public Sector: Selected Lessons of International Experience,* ECD Working Paper Serie, n°13, 2005

World Bank, *Evaluation Capacity Development, Institutionalization of Monitoring and Evaluation Systems to Improve Public Sector Management*, ECD Working Paper Series, n°15, IEG, janvier 2006

World Bank, *Comment mettre en place des systèmes de S&E pour améliorer les performances du secteur public*, 2007

World Bank, *Influencing Change: Building Evaluation Capacity to Strengthen Governance*, 2011

Travaux de recherche

BERTHELEMY, Jean-Claude, *Sur la déclaration de Paris, Évaluer l'évaluation,* document de travail 41, FERDI, mars 2012

BOURGUIGNON, François, FEREIRA, Francisco, "Ex Ante Evaluation of Policy Reforms Using Behavioral Models", from *Tool Kit for Evaluating the Poverty and Distributional Impact of Economic Policies*, World Bank, 2003

BIRDSALL, Nancy, KHARAS, Homi, *The quality of official development assistance, (QUODA)*, Third Edition, CGDev, Brookings Institute, 2014

CGDev, *When will we ever learn? Improving lives through impact evaluation*, Report of the evaluation gap working group, mai 2006

CLEAR, *Demand and Supply: Monitoring, Evaluation, and Performance Management Information and Service in Anglophone Sub-Saharan Africa*, 2013

CLEAR-CESAG, *An Assessment of Demand for Monitoring, Evaluation, and Performance Management Information and Services*, Prepared for the CLEAR Francophone Africa Center at the Centre africain d'études supérieures en gestion, 2013

DEATON, Angus, *Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development",* Manuscript, Princeton, janvier 2009

EYBEN, Rosalind, *Uncovering the politics of evidence and results, a framing paper for development practitioners*, Big Push Forward, avril 2013

Institute for Development, Studies, *Participatory Monitoring and Evaluation: An Introductory Pack*, University of Sussex, 1997

LOISEAU, François, « Les pratiques d'évaluation dans les organisations internationales », *Cahier du CEVIPOF*, 2006

Mutual Organizations Performance Assessment Network, *Annual Report*, 2013

NDIKUMANA, Leonce, *Applying evaluation to development and aid: can evaluation bridge the micro-macro gaps in aid effectiveness*, University of Massachusetts, Department of Economics and Political Economy Research Institute, 2012

ODI, *Improving impact evaluation production and use*, Working Paper, mars 2009

RODRIK, Dani, *The new development economics: we shall experiment, but how shall we learn?* Paper prepared for the Brookings Development Conference, 29-30 mai 2008

ROODMAN, David, *An Index of Donor Performance*, Working Paper n°67, CGDev, novembre 2006

SCHAUMBURG-MULLER, Henrik, *Evaluation Capacity Building, Donor Support and Experiences*, Report for the DAC Expert Group on Aid Evaluation, OCDE, 1997

SKOUFIAS, Emmanuel, *Progresa and its Impacts on the Welfare of Rural Households in Mexico*, Research Report n°139, International Food Policy Research Institute, 2005

EVALUATION DES POLITIQUES PUBLIQUES

Ouvrages

ALKIN, Marvin, *Debates on Evaluation*, Sage Publication, 1990, 304p

ARNAUD, Serge, BOUDEVILLE, Nicolas, *Évaluer des politiques et programmes publics*, Paris, Éditions de la Performance, 2004, 136p

BASLE, Maurice, DUPUIS, Jérôme, LE GUYADER, Sylviane, *Évaluation, action publique territoriale et collective*, tome 2, L'Harmattan, 2002, 332p

CONNELL, James, *et al*. (Eds.), *New Approaches to Evaluating Community Initiatives*, Aspen Institute, 1995, 255p

del RAY, Angélique, *La Tyrannie de l'évaluation,* La Découverte, Collection Cahiers libres, 2013, 160p

ELIADIS, Pearl, FURUBO, Jean-Eric, JACOB, Steve, *Evaluation, Seeking Truth or Power ?*, Transaction Publishers, 2011, 358p

FURUBO, Jan-Eric, RIST, Ray C, SANDAHL, Rolf, *International Atlas of Evaluation*, Transaction Publishers. Cop. 2002, 471p

GUBA, Egon G, LINCOLN, Yvonna S, *Fourth Generation Evaluation*, SAGE Publications, 1989, 249p

GUBA, Egon G, LINCOLN, Yvonna S, "The countenances of fourth generation evaluation: description, judgment and negotiation" in PALUMBO, Dennis, (dir.), *The politics of program evaluation*, Sage, 1987, 309p

FOUQUET, Annie, MEASSON, Ludovic, *L'évaluation des politiques publiques en Europe : cultures et futurs*, L'Harmattan, 2009, 444p

HOUSE, Ernest, *Professional Evaluation: Social Impact and Political Consequence,* Sage Publication, 1993, 205p

JACOB, Steve, *Institutionnaliser l'évaluation des politiques publiques : étude comparée des dispositifs en Belgique, en France, en Suisse et au Pays-Bas,* PIE-Peter Lang, 2006, 271p

LAVIGNE-DELVILLE, Philippe, MATHIEU, Marilou, SALLAMNA, Nour-Eddine, *Les enquêtes participatives en débat : ambitions, pratiques, enjeux,* Coédition Gret/Karthala/Icra, 2000, 544p

PATTON, Michael Q, *Utilization focused evaluations,* Sage Publication, 1978, 688p

PERRET, Bernard, *L'évaluation des politiques publiques,* Repères, 2008, 120p

ROSSI, Peter H, *et al*., *Evaluation: A systemic approach,* septième édition, SAGE Publications, 2004, 380p

RUSS-EFT, Darlène, *et al*., *Evaluators Competencies, Standards for the Practice of Evaluation in Organizations,* John Wiley and Son, 2008, 238p

SCHUMAN, Edward A, *Evaluative Research: Principles and Practice in Public Service and Social Action Programs,* Russel Sage Foundation, 1967, 186p

SCRIVEN, Michael, "The Methodology of evaluation, in TYLER, Ralph, GAGNE, Robert, SCRIVEN, Michael, *Perspective of curriculum evaluation,* Rand McNally, Chicago, 1967, 102p

SCRIVEN, Michael, ROTH, Jane, *Evaluation thesaurus,* Edgepress, 1991, 408p

RIEPER, Olaf, TOULEMONDE, Jacques, *Politics and Practices of International Evaluation,* NJ:Transaction Publishers, 1997, 172p

VIVERET, Patrick, *L'évaluation des politiques et des actions publiques*, La Documentation française, 1989

L'évaluation des politiques publiques en Europe: cultures et futurs, Colloque de Strasbourg, L'Harmattan, 2008, 442p

Articles

ALKIN, Marvin, C. "Naturalistic Study of Evaluation Utilization", *New Directions for Program Evaluations: Utilization of Evaluative Information,* vol. 5, 1980, pp. 19-28

BARON, Gaëlle, MONNIER, Éric, « Une approche pluraliste et participative: coproduire l'évaluation avec la société civile », *Revue informations sociales,* n°110, septembre 2003, pp. 1-21

BOURGEOIS, Isabelle, COUSINS, Bradley J, “Understanding Dimensions of Organizational Evaluation Capacity”, *American Journal of Evaluation,* vol. 34, n°3, 2013, pp. 299-319

BRUNNER, Lise, GUZMAN, Alba “Participatory evaluation: a tool to assess projects and empower people”, *New Direction for Program Evaluation,* Summer 1989, pp. 9-18

DEWEY, Jennifer D, “Evaluators’ Competencies, What’s Taught Versus What Sought’s”, *American Journal of Evaluation,* vol. 29, n°3, septembre 2008, pp. 268-287

DURAN, Patrice, MONNIER, Eric, SMITH Andy, « Evaluation à la française, Toward a New Relationship between Social Science and Public Action », *Evaluation,* vol. 1, n°1, juillet 1995, pp. 45-53

FAURE, Alain, SMITH, Andy, « L’évaluation objet de savoir et enjeu de pouvoir », *Sciences de la société,* n°32, pp. 149-152

GORI, Roland, del VOLGO, Marie-José, « l’idéologie de l’évaluation : un nouveau dispositif de servitude volontaire », *Nouvelle revue de psychologie,* vol. 2, n°8, 2009, pp. 11-26

JACOB, Steve, VARONE, Steve, « Institutionnalisation de l’évaluation et nouvelle gestion publique : un état des lieux comparatif », *Revue internationale de politique comparée,* vol. 11, février 2004, pp. 271-292

JACOBS, Steve, BOISVERT, Yves, “To be or not to be a profession: pros, cons and challenges for evaluation”, *Evaluation,* vol. 16, n°4, 2010, pp. 349-369

JAHODA, Marie, BARITZ, Edward, “The nature of evaluation”, *International Social Science Bulletin,* vol. 7, n°3, 1955, pp. 1-13

LINCOLN, Yvonna, “The ERS standards for program evaluation”, *Evaluation and Program Planning,* vol. 8, n°3, 1985, pp. 251-263

MATYJASIK, Nicolas, « Des évaluations de politiques et des consultants: émergence d’un champ professionnel et segmentation des rôles », *Revue française d’administration publique,* n°148, 2010, pp. 907-921

PATTON, Michael, “The challenge of being a Profession”, *Evaluation Practice,* vol. 11, n°1, 1990, pp. 45-51

PICCIOTTIO, Robert, “The logic of evaluation professionalism”, *Evaluation,* vol. 17, n°2, 2011, pp. 165-180

PRITCHETT, Lant, “It pays to be ignorant: A simple political economy of rigorous program evaluation”, *The journal of policy reform,* 2002, pp. 251-269

SCHAUMBURG-MULLER, Henrik, “Use of aid evaluation from an organizational Perspective”, *Evaluation,* vol. 11, n°2, 2005, pp. 207-222

SCRIVEN, Michael, “Types of Evaluators and Type of Evaluations”, *American Journal of Evaluatio*n, vol. 17, n°2, 1996, pp. 61-71

SIMONET, François, « L'évaluation : objet de standardisation des pratiques sociales », *Cités*, n°37, Janvier 2009, pp. 91-100

STEVAHN, Laurie, et *al*., "Establishing Essential Competencies for Program Evaluators", *Evaluation,* vol. 26, n°1, mars 2005, pp. 43-59

THOENIG, Jean-Claude, « l'évaluation source de connaissance applicable aux réformes de la gestion politique », *Revue française d'administration publique,* n°92, 1999, pp. 80-92

VALOVIRTA, Voimavirta, "Evaluation Utilization as Argumentation", *Evaluation,* vol. 8, n°1, 2002, pp. 60-80

WEISS, Carol H, "The stakeholder approach to evaluation: Origins and promise", *New Directions in Program Evaluation,* n°17, 1983, pp. 3-14

WEISS, Carole, "Have we learned anything new about the use of evaluation", *American Journal of Evaluation,* vol. 19, n°1, 1998, pp. 121-133

ZARKA, Yves-Charles, « l'Évaluation, un pouvoir supposé savoir », *Cités,* n°37, janvier 2009, 12p

Littérature grise

Conseil scientifique de l'évaluation, *L'évaluation en développement,* Rapport annuel sur l'évolution des pratiques d'évaluation des politiques publiques, La Documentation française, 1997

HM Treasury, *The « Green Book » : appraisal and evaluation in Central Government: Treasury guidance,* 2003

NAO, *Measuring the Performance of Government Departments,* 22 mars 2001

Société française d'évaluation, *Charte de l'évaluation des politiques publiques et des programmes publiques,* 28 octobre 2003.

TALBOT, Colin, *Performance in Government : The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom,* Evaluation Capacity Development, IEG World Bank, novembre 2010

The Better Regulation Executive, The Department of Work and Pensions, Research Methods on Policy Evaluation, 2001

Papiers de travail

CAMPBELL, Donald, *Assessing the impact of planned social change*, paper 8, occasional paper serie, University of Darthmouth, décembre 1976

FOURNIER, Marcel, GINGRAS, Yves, MATHURIN, Creutzer, « l'Évaluation par les pairs et la recherche légitime », *Actes de la Recherche en sciences sociales,* n°19, septembre 1988

LACOUETTE-FOUGERE, Clément, LASCOUMES, Pierre, *Les scènes multiples de l'évaluation : les problèmes récurrents de son institutionnalisation,* LIEPP, mai 2013

L'HORTY, Yannick, PETIT, Pascale, « L'évaluation aléatoire : un succès qui ne doit rien au hasard », *La vie des idées,* 11 mai 2011

WHITMORE, Elizabeth, "Understanding and Practicing Participatory Evaluation", *New Directions for Evaluation*, n° 80. Wiley,1998

Thèses et mémoires

MATYJASIK, Nicolas*, L'´évaluation des politiques publiques dans une France décentralisée. Institutions, marché et professionnels*, Thèse pour le Doctorat en Science politique sous la direction de Andy Smith, Université de Bordeaux, 3 décembre 2010, 405p

SPENLEHAUER, Vincent, *L'évaluation des politiques publiques, avatar de la planification,* Thèse pour l'obtention du doctorat de science politique de l'université Pierre Mendes France de Grenoble, 8 décembre 1998, 356p

SOCIOLOGIE POLITIQUE / SOCIOLOGIE DES ORGANISATIONS

Ouvrages

ARGYRIS, Chris, SCHÖN, Donald A, *Organizational learning II: Theory, Method and Practice.* Reading, MA: Addison-Wesley, 1996, 305p

AXELROD, Robert, T*he Evolution of Cooperation*, Basic Books, 1984, 264p

BARZELAY, Michael, *The New Public Management : improving research and policy dialogue,* Russel Sage Foundation, 2000, 218p

BECK, Ulrich, *La société du risque : sur la voie d'une autre modernité,* Flammarion, 2008, 521p

BRECHIN, Steven, *Planting Trees in the Developing World: a sociology of international organizations,* The Johns Hopkins University Press, 1997, 280p

BAUDOUIN, Jean, *Introduction à la sociologie politique,* Éditions du Seuil, 1998, 324p

CARTER, Nail, KLEIN, Patricia, DAY, Rudolf, *How Organizations Measure Success: the use of performance indicators in Government*, Routledge, 1992, 208p

COMTE, Auguste, *Discours sur l'ensemble du positivisme*, Flammarion, ed. 1999, 1848, 464p

CROZIER, Michel, *Le système bureaucratique,* Éditions du Seuil, 1963, 369p

CROZIER, Michel, FRIEDBERG, Erhard, *L'acteur et le système : les contraintes de l'action collective,* Éditions du Seuil, 1981, 500p

DENNING, Stephen, *The Springboard, How Storytelling Ignites Action in Knowledge-Era Organizations,* KMCI Press, Routledge, 2000, 248p

DEROSIERES, Alain, *La politique des grands nombres, essai sur la raison statistique*, La Découverte, 2000, 456p

DURKHEIM, Émile, *Les règles de la méthode sociologique*, les Presses universitaires de France, 16e ed., 1967, première publication en 1894, 149p

FERLIE, Ewan, *The New Public Management in Action,* Oxford University Press, 1996, 287p

FOUCAULT(b), Michel, *et al*., *Naissance de la biopolitique : cours au Collège de France (1978-1979),* Gallimard : Seuil, 2004, 355p

FOUCAULT(c), Michel, *Du Gouvernement des vivants : cours au Collège de France (1979-1980),* Gallimard : Seuil, 2004, 380p

FOUCAULT, Michel, *Sécurité, territoire, population : cours au Collège de France (1977-1978),* Gallimard : Seuil, 2004, 435p

FREEMAN, Edward R, *Strategic Management: a Stakeholder Approach*, Cambridge University Press, 2010, Première publication 1984, 292p

HIBOU, Béatrice, *La bureaucratisation du monde à l'ère néolibérale,* Éditions la Découverte, Paris, 2012, 223p

HOMANS, George, *Social Behavior: Its Elementary Forms,* Harcourt, Brace & World, 1961, 441p

KINGDON, John W, *Agendas, alternatives, and public policies,* Longman, cop, 2nd ed., 2003, 253p

LASCOUMES, Pierre, LE GALES, Patrick, *Gouverner par les instruments*, Presses de Sciences po, 2005, 369p

MACDONALD, Kate M, *The Sociology of the Professions,* Thousand Oaks, CA: SAGE, 1995, 224p

MAUSS, Marcel, *Essai sur le don, forme et raison de l'échange dans les sociétés archaïques,* Texte extrait de l'Année sociologique, seconde série, 1924-1925, Éditions PUF, 2007, 248p

MILLERSON, Geoffroy, *The Qualifying Associations: A Study in Professionalization,* Routledge and Kegan Paul Ltd, 1964, 304p

OSBORNE, David, GAEBLER, Ted, *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector,* Plume, 1993, 432p.

POWELL, Walter W, Di MAGGIO, Paul J *The New Institutionalism in Organizational Analysis*, Chicago, University of Chicago Press, 1999, 486p

POWER, Michael, *Organized Uncertainty, Designing a world of risk Management*, Oxford University Press, 2007, 248p

SENGE, Peter, *The fifth discipline: the art and practice of the learning organization*, Doubleday/Currency, 1990, 424p

Articles

HOOD, Christopher, "A public management for all seasons", *Public Administration*, 69, spring, 1991, pp. 3-19

POLLIT, Summa, "Trajectories of Reform: Public Management Change in Four Countries", *Public Money and Management*, vol. 17, n°1, janvier-mars, pp. 1-13

Littérature grise

DIFD, *Doing the knowledge: how DFID compares with best practices in knowledge management*, 2000

HOVLAND, Ingie, *Knowledge Management and organizational learning: an international development perspective*, ODI, août 2003

KING, Kenneth, Mc GRATH, Simon, *Knowledge sharing in development agencies: lessons from four cases*, the World Bank operations evaluation department, 2003

ODI, *Strengthening learning from research and evaluation: going with the grain*, septembre 2010

SIDA, *knowledge and learning in aid organizations: a literature review with suggestions for further studies*, SADEV working paper, janvier 2007

Papiers de recherche

ALMOND Gabriel, VERBA, Sydney, *The Civic Culture, Political Attitudes and Democracy in Five nations*, Princeton University Press, 1963

AMAR, Anne, BERTHIER, Ludovic, *Le nouveau Management Public : Avantages et limites*, CEROG, 2006

POLITIQUES PUBLIQUES

Ouvrages

COMMON, Richard, *Public Management and Policy Transfer in Southeast Asia,* Aldershot, 2001, 300p

EVANS, Mark, *Policy Transfer in Global Perspective,* Aldershot, Ashgate, 2004, 250p

IKENBERRY, John G, "The International Spread of Privatization Policies : Inducements, Learning and Policy Bandwagoning", in SULIEMAN, Ezra, WATERBURY, John N, *The Political Economy of Public Sector Reform and Privatization,* Boulder, San Francisco, Oxford, Westview, 1990, 388p

WESTNEY ELEANOR, *Imitation and Innovation: the Transfer of Western Organizational Patterns to Meiji Japan,* Cambridge, Harvard University Press, 1987, 256p

Articles

CAMPBELL, Donald T, "Reforms as experiments", *American Psychologist,* n°24, 1969, pp. 409-429

PERRET, Bernard, « De l'échec de la rationalisation des choix budgétaires (RCB) à la loi organique relative aux Lois de Finances », *Revue française d'Administration publique,* n°117, 2006, pp. 31-41

WOLMAN, Harold, "Policy transfer among local governments: an information-theory approach", *Governance,* vol. 15, n°4, 2002, pp. 379-396

Rapports gouvernementaux

COLEMAN, James S, *Policy Analysis on Major Issues: a compilation of papers prepared for the Commission on the Operation of the Senate,* US Government Printing Office, Washington DC, 1972

GODELIER, Maurice, *Les Sciences de l'homme et de la société en Franc : analyse et propositions pour une politique nouvelle : rapport général,* Ministère de la recherche et de l'industrie, 1982

OLSON, Mancur, *Toward a social report,* US Department of Health, Education, and Welfare. Washington: U.S Government Printing Office, 1969

Prime Minister and the Minister for the Cabinet Office, *Modernising Government,* mars 1999

Travaux de recherche

TAPLIN, Dana, *et al*. *Theory of Change, a Serie of Papers to Support Development of Theories of Change based on Practice in the Field,* ActKnowledge, avril 2013

DELPEUCH, Thierry, « L'analyse des transferts internationaux de politique publique : un état de l'ar*t »*, *Questions de recherche*, n°27, CERI, décembre 2008

SITES INTERNET

The Broker http://www.thebrokeronline.eu/

Brookings http://www.brookings.edu/

Center on Foreign Relations http://www.cfr.org/

CGdev
http://www.cgdev.org/section/publications?f%5B0%5D=field_document_type%3A2023

CLEAR http://www.theclearinitiative.org/

COMPAS http://www.mfdr.org/COMPAS/documents/2010_COMPAS-Report.pdf

DAC-OECD http://www.oecd.org/fr/cad/

DAC-OECDEvalnet
http://www.oecd.org/dac/evaluation/evaluatingmultilateraleffectiveness.htm

DAC-OECD evaluation resource center http://www.oecd.org/derec/

DAC-OECD evaluation policies and guidelines http://www.oecd.org/derec/guidelines.htm

DEVEX https://www.devex.com/en/

Development Impact initiative (World Bank)
http://web.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTDEVIMPEVAINI/0,,contentMDK:23422972~pagePK:64168445~piPK:64168309~theSitePK:3998212,00.html

Development Gateway http://www.developmentgateway.org/

DEVPOLICY http://devpolicy.org/busan-a-new-global-divide-on-aid20111202/

Evaluation Capacity Development Group http://www.ecdg.net/

GLOBAL Fund http://www.theglobalfund.org/fr/about/results/

Government Offices of Sweden http://www.government.se/sb/d/11747/a/122004

IDEAS http://www.ioce.net/en/index.php

IEG (Evaluation Capacity Development working papers) http://ieg.worldbankgroup.org/ieg-search?f[0]=field_document_categories%3A334&f[1]=field_doc_sub_category%3A215

IEG (Evaluation Capacity Development guides) http://ieg.worldbankgroup.org/ieg-search?f[0]=field_document_categories%3A334&f[1]=field_doc_sub_category%3A339

Initiative 3IE http://www.3ieimpact.org/

IOCE http://www.ioce.net/en/index.php

IPDET http://www.ipdet.org/

KOICA http://www.koica.go.kr/english/koica/policies/index.html

Ministry of Foreign Affairs, Republic of Korea, ODA/Development Cooperation page http://www.mofat.go.kr/ENG/policy/oda/index.jsp?menu=m_20_11

MOPAN http://www.mopanonline.org/

MYMANDE http://www.mymande.org/

Open Democracy http://www.opendemocracy.net/

People's Daily http://english.people.com.cn

Plateforme aide au développement Laos www.rtm.org.la

QWIDS (statistiques du CAD-OCDE) http://stats.oecd.org/qwids/