

L'évaluation de politique publique : un exemple de pratique d'amélioration des institutions publiques en tension entre efficience et pluralisme

Laurent Barbut, Marielle Berriet-Sollic, Thomas Péguy

► To cite this version:

Laurent Barbut, Marielle Berriet-Sollic, Thomas Péguy. L'évaluation de politique publique : un exemple de pratique d'amélioration des institutions publiques en tension entre efficience et pluralisme. Colloque international : "Performances et institutions : de l'efficience au pluralisme?", Université de Reims Champagne Ardenne (URCA). FRA., Nov 2011, Reims, France. 16 p. hal-02808782

HAL Id: hal-02808782

<https://hal.inrae.fr/hal-02808782>

Submitted on 6 Jun 2020

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.

Communication au colloque de Reims, « Performance et Institutions : de l'efficience au pluralisme », 17 et 18 novembre 2011

**L'évaluation de politique publique :
un exemple de pratique d'amélioration des institutions publiques
en tension entre efficience et pluralisme**

Laurent BARBUT¹ Marielle BERRIET-SOLLIEC ² & Thomas PEGUY³? &

Résumé

Démarche héritée des pratiques anglo-saxonnes, l'évaluation de politique s'est progressivement développée en France depuis les années 90. Dans la plupart des cas, l'enjeu pour le commanditaire est de « mesurer » et d'améliorer la performance des institutions publiques. La présente communication propose une analyse critique de ce développement de l'évaluation de politique publique depuis une vingtaine d'années et analyse en particulier la place et le rôle d'une part de la Commission Européenne, d'autre part de la Société Française de l'Evaluation dans cette évolution. Les auteurs, en tant qu'acteurs de ce développement à différents titres (consultant, administratif et enseignant-chercheur), proposent d'appuyer leurs analyses sur leurs expériences et pratiques relatives à l'évaluation des politiques interministérielles à l'échelle nationale, et à l'évaluation dans les pays membres de deux politiques majeures de la Commission Européenne : Politique Agricole Commune et politique dite de cohésion.

Mots clés : Evaluation, performance, action publique, PAC, politique de cohésion

¹ Directeur du bureau d'études EPICES, 46 rue Bichat, 75010 PARIS, et co-animateur du groupe de travail « Evaluation des politiques territoriales » au sein de la Société Française d'Evaluation (SFE). Mél : laurent.barbut@epices-net.fr

² Professeure d'économie à AgroSup Dijon, membre de l'unité de recherche CESAER (UMR 1041 INRA-AgroSupDijon)

³ Membre de l'équipe « Développement régional et politiques européennes » à la DATAR.

Introduction

L'évaluation de politique publique est une pratique d'amélioration de l'action publique originaire des pays de culture anglo-saxonne (Shadish et al., 1991 ; Rossi et Freeman, 1993 ; Patton, 2008) et en développement en France depuis le début des années 1990 (Nioche et Poinard, 1984 ; Monnier, 1987 ; Greffe, 1989 ; Perret, 2001). Elle a été institutionnalisée au sein de l'appareil d'Etat par le décret du 22 janvier 1990, qui la définit ainsi « Evaluer une politique, c'est rechercher si les moyens juridiques, administratifs ou financiers mis en œuvre permettent de produire les effets attendus de cette politique et d'atteindre les objectifs qui lui sont fixés ».

Ainsi, à travers l'analyse de l'efficacité des politiques publiques — résultats obtenus au regard des objectifs fixés — et de leur efficience — résultats obtenus au regard des moyens engagés — l'évaluation de politique publique entre dans le vaste champ des pratiques visant à « mesurer » et à améliorer la performance des institutions publiques. En ce sens, l'analyse de son développement en France et/ou au sein de la Commission Européenne, à travers les débats et controverses qui l'accompagne, est particulièrement riche pour éclairer la problématique de la prise en compte des différentes dimensions de la performance des institutions publiques. En effet, on retrouve dans l'analyse de l'évolution des doctrines et des pratiques certains débats récurrents comme :

- la définition même de l'évaluation de politiques publiques, avec des visions plus ou moins larges de ses finalités et des référentiels auxquels confronter les analyses⁴ ;
- la place à accorder dans les évaluations aux approches quantitatives par rapports aux approches qualitatives ;
- la place à accorder aux indicateurs de suivi quantifiés ;
- la temporalité de l'évaluation — souvent considérée comme longue — par rapport au calendrier de la décision politique ;
- le respect systématique ou non de la dimension pluraliste (prise en compte de la pluralité des points de vue) de l'évaluation.

Ces questions ne sont pas très éloignées de celles agitant les débats sur la mesure de la performance économique ou de l'utilité sociale, et renvoient dans le champ des institutions publiques aux tensions entre court terme et long terme et/ou entre efficacité budgétaire et démocratie.

Nature et démarche de la communication proposée

⁴ Citons par exemple, en comparaison de celle du décret qui vient d'être évoqué, la définition de Patrick Viveret dans son rapport au premier ministre de 1989 « Evaluer une politique, c'est former un jugement sur sa valeur ».

La communication proposée repose sur une analyse du développement de l'évaluation de politique publique depuis une vingtaine d'années au sein (1) de l'Etat français, (2) du réseau d'acteurs constitué en France autour de cette pratique au sein de la Société Française de l'Evaluation, et (3) des pratiques de la Commission Européenne. Les auteurs, en tant qu'acteurs de ce développement à différents titres (consultant, administratif et enseignant-chercheur), proposent d'appuyer leur analyse sur leurs expériences et pratiques. Cette analyse s'appuie plus particulièrement sur trois exemples : en premier lieu celui de l'évaluation des politiques interministérielles à l'échelle nationale, en second lieu celui des débats et controverses faisant l'actualité de la Société Française de l'Evaluation et en troisième lieu celui de l'évaluation dans les pays membres de deux politiques majeures de la Commission Européenne (Politique Agricole Commune et politique dite de cohésion). A l'issue de la présentation de ces trois regards sur le développement de l'évaluation de politique publique, nous en proposons une analyse croisée sous trois angles successifs, dont le troisième tiendra lieu de conclusion :

- les exemples développés mettent-ils en évidence une évolution des acteurs fortement impliqués dans l'évaluation au sein des institutions ?
- ces mêmes exemples soulignent-ils une évolution des doctrines et/ou des méthodes préconisées ?
- et enfin, comment ces exemples éclairent-ils le débat sur le pluralisme de la mesure de la performance des institutions, à travers le prisme d'un de ses instruments, l'évaluation de politique publique ?

1. Deux décennies de hauts et de bas pour l'évaluation interministérielle au sein de l'appareil d'Etat

Les prémisses de l'évaluation de politique publique en France peuvent être datées des années 1970 avec la Rationalisation des Choix Budgétaires (RCB), qui visait à restructurer le budget sous forme d'un ensemble de programmes d'action et à fonder les décisions de dépense sur une prévision de leurs impacts socio-économiques. Mais ce n'est qu'à la fin des années 1980, sous le gouvernement de Michel Rocard, que l'évaluation sera véritablement reconnue en France (Perret, 2001). Suite au rapport élaboré par Patrick Viveret, le décret du 22 janvier 1990 institutionnalise l'évaluation au sein de l'Etat français en créant le Comité interministériel de l'évaluation (CIME), le Conseil Scientifique de l'Evaluation (CSE) et le Fonds national de développement de l'évaluation (FNDE). Le Commissariat général du Plan est chargé de l'animation administrative du dispositif, et, de 1990 à 1998, une quinzaine d'évaluations interministérielles de grande ampleur sont menées à bien. Leur durée excessive et leurs faibles suites décisionnelles directes provoquent l'enlisement administratif de la procédure. Pour tenter de relancer cette évaluation interministérielle, un nouveau décret est pris le 18 novembre 1998, dans le but de simplifier et de dynamiser la procédure de l'évaluation interministérielle. Mais le nouveau dispositif butera comme le précédent sur les lourdeurs du travail interministériel, et sera progressivement abandonné au début des années 2000.

Au cours de cette décennie 1990/2000 où le Commissariat Général au Plan joue un rôle central, le rôle du Parlement est en revanche faible dans le développement de l'évaluation au sein de l'appareil d'Etat. Une première tentative émane en 1996 de Philippe Séguin (alors président de l'Assemblée nationale), qui fait voter deux lois instituant deux structures communes aux deux assemblées : un Office parlementaire d'évaluation des lois et un Office parlementaire d'évaluation des politiques publiques. Ces deux Offices seront rapidement mis en sommeil⁵ au profit d'une Mission d'évaluation et de contrôle (MEC) relevant uniquement de l'Assemblée nationale⁶, et formalisée juridiquement par la LOLF en 2001. La mise en place de cette Loi Organique relative aux Lois de Finances est alors au centre des préoccupations des Ministères et alimente largement les débats dans le microcosme des acteurs de l'évaluation (nous y reviendrons dans la partie suivante). L'évaluation interministérielle pour sa part a pour ainsi dire disparu en cette première partie de la décennie 2000.

Elle ne sera relancée qu'en 2007, avec la création de la Mission d'Evaluation des Politiques Publiques (MEPP), d'abord sous l'égide de l'éphémère Secrétariat d'Etat chargé de la Prospective et de l'Evaluation des politiques publiques (2007/2009), puis sous l'autorité du Ministre du Budget et de la Réforme de l'Etat. La MEPP est chargée de piloter les structures d'évaluation interne de l'administration : elle assure à l'échelle interministérielle la gouvernance de l'évaluation et la coordination des Comités Ministériel d'Evaluation (CME). Elle anime un Comité Interministériel d'Evaluation (CIE), chargé de la programmation d'évaluations interministérielles. Ces dernières sont néanmoins très peu diffusées aujourd'hui.

Par ailleurs, en parallèle de cette relance pilotée par le Ministère du Budget, une responsabilité en matière d'évaluation des politiques de l'Etat est confiée en 2008 à l'Assemblée nationale. En effet, la dernière révision constitutionnelle⁷, précise dans l'article 24 de la Constitution, que non seulement « Le Parlement vote la loi et contrôle l'action du Gouvernement » mais aussi qu' « il évalue les politiques publiques ». Pour assurer cette responsabilité accrue, il est apparu souhaitable de créer au sein de l'Assemblée un organe spécifique pour dépasser les limites des commissions permanentes dans le cadre d'évaluations de politiques à forte dimension transversale. C'est la raison de la création du Comité d'Evaluation et de Contrôle des politiques publiques de l'Assemblée nationale (CEC), dont la composition et les missions sont précisées dans les articles 146-2 à 146-7 du Règlement de l'Assemblée nationale. Afin d'assurer sa mission d'évaluation, le CEC et ses rapporteurs peuvent s'appuyer sur les services de l'Assemblée, ainsi que sur des experts extérieurs, comme notamment la Cour des comptes, chargée par l'article 47-2 de la Constitution d'une mission d'assistance au Parlement pour l'évaluation des politiques publiques. On note donc que le rôle du Parlement dans l'évaluation de politique publique au niveau national, de périphérique au début des années 1990 est censé devenir central à la fin de la décennie 2000.

Soulignons enfin, à l'issue de ce premier exemple, que ces aléas de l'évaluation interministérielle n'ont pas empêché l'évaluation de se développer parallèlement dans

⁵ Du fait de la difficulté à les faire fonctionner entre deux assemblées politiquement opposées.

⁶ Le Sénat créant de son côté son propre comité d'évaluation des politiques publiques le 15 juin 2000.

⁷ 23 juillet 2008.

d'autres cadres depuis le début des années 1990 : celui de comités thématiques ad hoc (Comité national d'évaluation de la recherche, comité national d'évaluation des universités, Haut Conseil de l'évaluation de l'école...), celui des ministères (créations d'unités spécialisées, services ou missions au sein des principaux ministères), celui des collectivités territoriales ou encore celui des établissements publics (comme la Caisse des Dépôts et Consignations par exemple).

Néanmoins, si ce développement de l'évaluation de politique publique dans les pratiques de l'action publique française est indéniable, il reste encore aujourd'hui confronté à de nombreuses réticences⁸, qui contribuent à diminuer la portée politique des évaluations réalisées (Berriet-Sollicec et al., 2005).

2. Une histoire de l'évaluation à travers les débats et controverses au sein de la Société Française d'Evaluation

La Société Française de l'Evaluation, créée en 1999 pour rassembler tous les acteurs parties prenantes de l'évaluation (administrations, chercheurs, prestataires du secteur privé), et dont le nombre d'adhérents augmente régulièrement depuis cette date, témoigne de la montée en puissance de l'évaluation des politiques publiques en France depuis deux décennies.

La SFE a pour vocation générale de contribuer au développement de l'évaluation et de promouvoir son utilisation dans la conception et la mise en œuvre des actions publiques et de toute action entreprise au nom de l'intérêt général. C'est un lieu de rassemblement, de débats, d'échanges, de capitalisation et de formation entre pairs, ouvert à tous les acteurs de l'évaluation : commanditaires, praticiens, utilisateurs, consultants, chercheurs et opérateurs privés (associations notamment) chargés de la mise en œuvre des politiques évaluées. La SFE s'attache tout particulièrement à :

- affirmer l'évaluation comme une démarche volontaire et nécessaire au sein des administrations et établissements publics, des collectivités territoriales, ainsi que dans les fondations et associations œuvrant pour l'intérêt général ;
- valoriser l'apport de l'évaluation à la conception, au pilotage et à la mise en œuvre des politiques publiques ;
- faciliter les échanges de "bonnes pratiques" ;
- accompagner la montée en compétence des acteurs de l'évaluation ;
- rendre intelligible l'évaluation auprès de tous - décideurs, gestionnaires, usagers, citoyens - et œuvrer pour la reconnaissance de la valeur qu'elle ajoute aux métiers du contrôle, de l'audit, du pilotage, de l'inspection et aux autres activités voisines.

⁸ Culture administrative encore peu favorable, poids de l'expertise dans la culture étatique française, confusion fréquente avec le contrôle, méfiance du politique...

Afin d'assurer ses missions de promotion de l'évaluation, de mise en réseau et de capitalisation, la SFE s'appuie sur un certain nombre d'activités, dont les trois principales sont :

- des réunions régulières de clubs régionaux et de groupes thématiques ;
- l'organisation d'un colloque annuel (jusqu'en 2004) puis bisannuel, réunissant plusieurs centaines de participants sur deux jours ;
- la publication d'ouvrages (actes des colloques, cahiers thématiques).

Les thèmes abordés dans les colloques ou les publications thématiques (voir tableaux 1 et 2 ci-après) fournissent une bonne illustration des débats qui ont traversé la SFE au cours de son histoire, elle même concomitante de deux décennies de développement de l'évaluation en France.

Tableau 1 : Thèmes des colloques de la SFE

Année	Thème
2001	Evaluation, action publique territoriale et collectivités
2002	L'évaluation au service de la stratégie publique ?
2003	La dynamique de l'évaluation face au développement durable
2004	Evaluation des politiques sociales
2006	Le citoyen, l'élu, l'expert : pour une démarche pluraliste d'évaluation des politiques publiques
2008	L'évaluation des politiques publiques en Europe : cultures et futurs
2009	Les nouvelles frontières de l'évaluation – 1989/2009 : vingt ans d'évaluation de politiques publiques en France, et demain ?
2011	L'évaluation des politiques publiques devant les défis d'une société en tension

Tableau 2 : Intitulés des cahiers de la SFE

Année	Thème
2005	Charte de l'évaluation : témoignages, débats et mise en œuvre des standards
2006	Evaluation et mesure de la performance
2007	La participation des citoyens à l'évaluation des politiques publiques
2009	Voisinages : de la confusion à la complémentarité — évaluation, audit de performance, management par la qualité
2010	Statistique et évaluation : quelle valeur (appréciation) donner à une valeur (mesure) ?

Au début des années 2000, la SFE structure ses activités et accompagne la montée en puissance d'une activité dont les acteurs se professionnalisent peu à peu. Sa première production significative, après d'importants débats, est celle d'une **charte de l'évaluation** (2005) qui pose les principes déontologiques de cette pratique et qui fait toujours largement référence. Au cours de la deuxième partie des années 2000, les débats internes vont se structurer autour de trois grandes thématiques, plus ou moins prégnantes selon les années et le contexte socio-économique.

La première est la définition du champ de ce qui relève de l'évaluation de politique publique, et de ce qui n'en relève pas, au regard des nombreuses autres pratiques d'amélioration de l'action publique elles aussi en plein développement : audit de performance, contrôle de gestion, management par la qualité,... Ce besoin de clarification apparaît notamment avec l'adoption en 2001 et la mise en application à partir de 2006 de la Loi Organique relative aux Lois de Finances (LOLF). Voulu comme une véritable révolution, la LOLF entend introduire la logique de performance dans la gestion publique et faire passer une culture budgétaire orientée vers les moyens à une logique de résultats. Elle vise également à introduire davantage de transparence dans l'information budgétaire afin d'en faciliter le contrôle par le Parlement. La loi organique instaure des Projets Annuels de Performance (PAP), qui présentent les actions des différentes administrations pour l'année à venir, avec des objectifs à atteindre. L'évaluation des objectifs se fait l'année suivante, dans les Rapports Annuels de Performance (RAP), sur la base d'un certain nombre d'indicateurs. La LOLF peut-elle sur ces bases se réclamer de l'évaluation de politique publique, comme le soutient nombre de ses promoteurs ? Le débat a été particulièrement vif sur cette question dans le monde de l'évaluation au milieu des années 2000. Il renvoie plus largement à la question des frontières entre évaluation, audit de performance et contrôle de gestion, qui est loin d'être clos, comme en témoigne le thème d'un cahier publié par la SFE en 2009 (tableau 2).

La deuxième thématique importante qui structure les débats est celle du principe de pluralité⁹, qui traite des acteurs à associer à l'évaluation d'une politique publique. Doit-on se limiter aux acteurs gestionnaires (élus, techniciens) ou associer les bénéficiaires et/ou les usagers, voire les citoyens à la formulation d'un jugement sur une politique publique visant à l'améliorer ? Cette question, on le comprend aisément, soulève à la fois des enjeux d'ordre éthique (autour du fonctionnement de la démocratie) et des enjeux d'ordre méthodologique (comment organiser en pratique une telle association ?). Ces enjeux seront le thème d'un colloque (2006) et d'un cahier thématique de la SFE (2007), et alimenteront un débat interne sur la rédaction du principe de pluralité dans la charte de l'évaluation (2011).

La troisième et dernière thématique importante au centre des débats est celle de la mesure des effets attribuables à une politique publique¹⁰, qui renvoie plus largement à la portée des argumentaires mobilisés dans les évaluations. Une telle mesure, de façon quantifiée et statistiquement robuste, est extrêmement difficile à réaliser et nécessite des méthodes d'investigation lourdes qui ne peuvent être mises en œuvre que dans un nombre réduit de situations. Le succès rencontré aujourd'hui, auprès des scientifiques comme des décideurs publics, par les démarches économiques expérimentales (Duflo et Kremer, 2005) atteste de cet engouement pour les techniques quantitatives. Dès lors, le débat se pose dans les termes suivants : les seules évaluations validées scientifiquement par des tests statistiques aléatoires (cf : essais contrôlés randomisés) doivent-elles se limiter aux cas où cette mesure est possible, ou bien les approches plus qualitatives, qui permettent d'une part d'approcher les effets là où leur mesure quantifiée est impossible et d'autre part de répondre à de nombreuses autres questions sur les politiques

⁹ Premier des sept principes énoncés dans la charte de la SFE précédemment citée.

¹⁰ Question qui intéresse souvent de façon prioritaire le commanditaire politique de l'évaluation.

publiques¹¹, sont-elles tout aussi utiles et légitimes ? Cette controverse est notamment au centre du dernier cahier publié par la SFE (2010). Elle alimente également les débats scientifiques sur la place des démarches empiriques et sur la validité des connaissances mobilisées en évaluation de politique (Berriet-Sollicet et al., 2011)

Soulignons pour terminer l'exposé de ce deuxième regard sur le développement de l'évaluation en France que ces débats et controverses doivent être examinés à l'aune du contexte général entourant l'action publique dans la période considérée, qui se caractérise par :

- d'une part, une tension de plus en plus forte sur les finances publiques, en particulier au cours de la décennie 2000, marquée par une faible croissance en Europe et des crises financières importantes (2008, 2011) ;
- d'autre part, une perte de confiance et de légitimité vis-à-vis des acteurs de la politique et de l'action publique, à la fois soupçonnés d'incapacité face à la puissance des acteurs financiers et traversés par de nombreux scandales renvoyant peu ou prou à la question de la transparence vis-à-vis de l'exercice du pouvoir confié par le vote.

3. L'évaluation dans les Etats membres de deux politiques européennes majeures

L'évaluation de politique publique étant une pratique développée à l'origine dans les pays de culture anglo-saxonne, c'est à travers eux qu'elle a précocement été introduite dans les pratiques de la Commission Européenne et synthétisées dans des documents méthodologiques, les dossiers MEANS, largement diffusés dans les Etats-membres (CE, 1999). Nous nous intéressons dans ce troisième regard à tout ou partie de deux politiques majeures dans le budget de l'Union Européenne : la Politique Agricole Commune (PAC) d'une part, et la politique dite de cohésion, qui vise à réduire les inégalités de développement au sein des Etats membres, d'autre part. Une partie de la PAC et la totalité de la politique de cohésion font l'objet d'un cofinancement entre l'Union Européenne et chaque Etat membre bénéficiaire. De ce fait, les objectifs poursuivis et les mesures mises en œuvre font l'objet de programmes pluriannuels négociés entre la Commission Européenne et tout ou partie de chaque Etat-membre¹². Chacun de ces programmes est ensuite piloté par une autorité de gestion nationale ou régionale, et suivi par un comité associant l'ensemble des acteurs impliqués (y compris la Commission Européenne). Tous les programmes sont soumis à une obligation de suivi (rapports annuels d'exécution) et d'évaluation à différents moments clés de leur mise en œuvre (à mi-parcours et lors de la dernière année du programme, par exemple), et c'est l'autorité de gestion de chaque programme qui a la responsabilité de mener ces évaluations et de les transmettre à la Commission Européenne. On constate donc à travers cette rapide présentation que ces évaluations se caractérisent par une « double commande » : celle de l'autorité de gestion du programme, qui a légitimement ses propres interrogations, et celle de la Commission Européenne, cofinanceur des programmes et destinataire final des évaluations réalisées.

¹¹ La politique répond-elle aux besoins de la société ? Comment expliquer qu'elle n'ait pas eu l'effet escompté ? Pourquoi tel effet non prévu s'est-il manifesté ?...

¹² Certains programmes sont régionaux, ou pluri-régionaux.

L'évolution depuis le début des années 1990 de l'évaluation de ces programmes relevant de la PAC ou de la politique de cohésion est marquée par quelques traits saillants.

Le premier d'entre eux est une convergence progressive des approches entre les deux Directions Générales en charge de ces deux politiques (DG Agri pour la PAC et DG Régio pour la politique de cohésion). Au début des années 1990, en effet, la culture de l'évaluation est plus ancienne au sein de la politique régionale, organisée depuis longtemps sous forme de programmes pluriannuels cofinancés alors que cette forme de mise en œuvre débute seulement pour la PAC (Daucé, 1998). Au début des années 2000, la culture de l'évaluation est présente pour les deux politiques, mais l'approche de la DG Agri est globalement beaucoup plus « cadrée » vis-à-vis des autorités de gestion alors que celle de la DG Régio laisse plus de liberté dans la conduite de chaque évaluation. Enfin, à la fin des années 2000, le dernier rapport sur la politique de cohésion prône une approche plus stratégique, davantage axée sur les résultats. Cette vision est très proche de celle que l'on peut lire dans la communication de la Commission sur la PAC à l'horizon 2020 : « Pour que les objectifs politiques puissent se traduire par des résultats sur le terrain, il est primordial que les mécanismes de mise en œuvre soient efficaces. L'approche stratégique actuelle serait renforcée au moyen d'objectifs quantifiés fixés au niveau de l'UE puis au niveau des programmes... Une telle évolution vers une approche davantage axée sur les résultats orienterait idéalement la politique vers les priorités de l'UE et mettrait en évidence ses résultats concrets ».

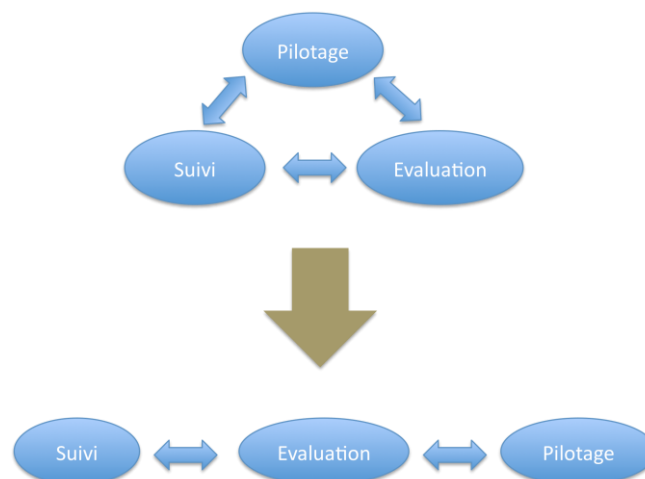
Le deuxième trait saillant de l'évolution de la doctrine en matière d'évaluation des programmes cofinancés de la PAC ou de la politique de cohésion est la recherche d'une meilleure articulation entre l'évaluation et le pilotage des programmes. L'objectif est ici de rendre plus opérationnelles les pratiques évaluatives et de valoriser la fonction instrumentale de l'évaluation (aide à la décision), telle qu'elle a pu être définie aux Etats-Unis, dès les années 80, par M. Patton autour de la notion d'« utilisation focalised evaluation » (2008). L'analyse des services de la Commission est en effet que des évaluations de qualité sont souvent réalisées, mais que leur utilisation par les acteurs pour faire évoluer les objectifs ou la mise en œuvre des programmes reste très limitée. Cette recherche de « connexion » entre évaluation et pilotage passera par différents vecteurs au cours du temps. Au début des années 2000, tout d'abord, les programmes de mise en œuvre de la politique de cohésion sont soumis à une réserve de performance. Le principe en est le suivant : une partie des fonds négociés en début de période est « mis en réserve » par la Commission, et est attribuée en cours de programme en fonction de critères de bonne gestion (avancement financier, taux de renseignement des indicateurs, atteinte des objectifs quantifiés...) et de la qualité des évaluations, dont l'importance est par ce biais renforcée. A partir de 2007, ensuite, pour l'ensemble des programmes cofinancés dans le cadre des deux politiques, la Commission prône l'évaluation dite « in itinere »¹³. Le principe en est que l'évaluation n'est plus nécessairement réalisée à des moments « fixes » du programme, mais lorsqu'un « problème » (risque de non atteinte des objectifs principalement) est détecté. L'évaluation est ainsi vue comme le moyen de fournir au(x) pilote(s) la connaissance nécessaire à la modification de la trajectoire. Ce type de pratique de l'évaluation est soumis à deux conditions importantes pour fonctionner : l'existence d'objectifs de résultats quantifiés, dont l'atteinte peut être

¹³ On retrouve ici la convergence de vue évoquée précédemment.

suivie dans le temps, et l'existence d'un système de suivi des programmes performant, permettant de suivre quasiment « en temps réel » l'atteinte des objectifs précédemment évoqués.

Enfin, le troisième trait saillant de cette évolution de doctrine évaluative découle des deux précédents : il s'agit de l'importance prise par les indicateurs. En effet le pilotage par les résultats nécessite à la fois un système de suivi performant et des indicateurs permettant de se situer par rapport à l'atteinte des objectifs fixés. Les indicateurs permettent par ailleurs des comparaisons entre des programmes devenus de plus en plus nombreux au fur et à mesure de l'élargissement des membres de l'Union Européenne. C'est tout d'abord dans les lignes directrices de la DG Agri qu'apparaissent, dès le début des années 2000, des indicateurs proposés par la Commission et devant être instruits de façon homogène dans l'ensemble des programmes. Cette approche est adoptée ensuite progressivement par la DG Régio, à la culture moins « centralisatrice » initialement. A partir de 2007, des indicateurs de réalisations et de résultats sont à fournir chaque année dans les rapports annuels de suivi des programmes cofinancés de la PAC, et ils occupent une place centrale dans les travaux d'évaluation à réaliser. La politique de cohésion devrait suivre la même voie puisque le dernier rapport sur cette politique stipule que « la présentation exhaustive, dans les délais impartis, d'informations précises sur les indicateurs et les progrès enregistrés dans la poursuite des objectifs définis serait capitale pour les rapports annuels ».

Ces trois éléments sont en fait liés entre eux et soulignent une évolution vers une doctrine unifiée entre les différentes directions de la commission. Cette évolution se caractérise par le passage d'une situation où l'évaluation était une activité particulière de la vie des programmes, liée à l'activité de suivi¹⁴ mais possédant néanmoins sa logique et son approche spécifique¹⁵, à une approche plus « linéaire » où l'évaluation est vue comme un prolongement de l'activité de suivi, servant à faire le lien entre cette dernière et la prise de décision opérationnelle. Le schéma ci-après résume ce passage d'une approche systémique de l'activité de suivi / évaluation à une approche linéaire



¹⁴ Parce qu'elle en valorise les données, notamment.

¹⁵ Pour approcher les effets et impacts, ou éclairer les enjeux de pertinence et de cohérence des programmes, à l'aide d'approches spécifiques (enquêtes, entretiens, études de cas,...).

Soulignons enfin, comme pour l'exemple précédent, que ces évolutions de la doctrine de la Commission Européenne ne peuvent être analysées sans rappeler deux éléments majeurs de contexte : la forte augmentation du nombre des pays membres, et les enjeux de gouvernance afférents, d'une part, et les fortes tensions budgétaires se répercutant des Etats membres sur l'Union Européenne, d'autre part.

4. Les acteurs clés de l'évaluation : une prise en main par les gestionnaires financiers ?

Le premier angle d'analyse transversal aux trois regards que nous venons de porter sur le développement de l'évaluation de politique publique nous conduit à nous intéresser aux acteurs clés des processus, c'est-à-dire ceux qui pilotent les dispositifs et/ou sont à l'origine des démarches d'évaluation, et ceux qui, de ce fait, en orientent la doctrine et les lignes directrices sur le plan méthodologique.

L'exemple de l'évaluation interministérielle, tout d'abord, souligne une nette évolution des acteurs en charge du pilotage du dispositif d'évaluation. En effet, alors que le Commissariat Général au Plan s'était vu confier cette mission au début des années 1990, c'est le Ministère du Budget qui en a la charge depuis la fin des années 2000. Par ailleurs ce premier exemple souligne également une volonté qui s'affirme en fin de période : celle de faire jouer un rôle central dans l'évaluation interministérielle au Parlement et à la Cour des Comptes.

Nous avons vu ensuite que la question de savoir quelles étaient les limites du champ de l'évaluation au sein de l'ensemble des pratiques d'amélioration de la gestion publique était au centre des débats internes aux membres de la SFE. Au delà des enjeux d'affirmation de son « territoire », la présence de cette question s'explique aussi largement par la nature des acteurs composant la SFE :

a) Au sein des administrations et collectivités territoriales commanditaires d'évaluation, tout d'abord, les membres adhérant à la SFE pilotent et/ou appartiennent à des directions ou à des services regroupant diverses missions d'appui transversal à la conduite de l'action publique (études, prospective, audit, contrôle de gestion, ...). Ils se révèlent souvent porteurs de ce débat sur les limites du champ de l'évaluation, qui les interpellent au gré de l'évolution du « rangement » de l'évaluation dans les organigrammes des services de l'Etat ou des collectivités. Or cette évolution semble marquée, au cours de la période que nous analysons, par un double mouvement : (1) un déplacement du rattachement de l'évaluation à des activités relevant des études, de l'observation, des statistiques, voire de la prospective vers des activités relevant du contrôle de gestion et/ou de l'audit interne, et (2) une remontée de la place de l'évaluation dans les organigrammes, avec un rattachement souvent proche de la Direction générale.

b) Au sein des consultants évaluateurs, ensuite, la montée en puissance de l'évaluation se traduit par une diversification de la taille économique des structures et l'arrivée sur le marché de l'évaluation des grands cabinets internationaux dont l'activité principale est l'audit financier et/ou la certification des comptes. Ces derniers portent également le débat sur les limites du champ de l'évaluation, qui n'est pas sans enjeu sur leur légitimité à exercer cette activité.

Enfin, en ce qui concerne l'apparition d'une doctrine unifiée de « pilotage par les résultats » au sein de différentes Directions Générales de la Commission Européenne, rien n'interdit de penser que la Direction Générale du Budget, dans un contexte particulièrement tendu, est au moins pour partie à l'origine de cette évolution des pratiques évaluatives.

Sur cette question des acteurs clés, les trois regards que nous avons porté sur l'évolution de la pratique évaluative semblent donc converger pour souligner une forme de « prise en main » de l'évaluation par les acteurs en charge de la gestion financière des institutions.

5. Les doctrines et les méthodes : plus de chiffres, moins de points de vue et controverses pris en compte ?

Le deuxième angle d'analyse transversal auquel nous nous intéressons est celui des doctrines et méthodes utilisées dans les processus d'évaluation, point sur lequel l'influence des acteurs clés précédemment évoqués n'est évidemment pas négligeable, mais qui renvoie également à d'autres influences comme les travaux de recherche ou les tendances observées dans d'autres pays. En toile de fond, cette interrogation soulève également la question des finalités assignées aux démarches évaluatives : s'agit-il de mesurer ? de comprendre ? de former les acteurs, parties prenantes des politiques (logique d'apprentissage) ? Force est de constater que selon les objectifs visés et la priorité retenue, les démarches et outils peuvent être radicalement différents (Berriet-Sollic et al., 2011).

Concernant tout d'abord l'évaluation interministérielle, nous avons souligné que le dispositif mis en place dans les années 1990 avait été abandonné, suite notamment à des attaques répétées sur sa lourdeur qui se traduisait par des durées de réalisation des évaluations jugées excessives (jusqu'à plusieurs années). Si ces durées nuisent indubitablement à l'utilisation politique des travaux d'évaluation, il convient de souligner qu'elles sont dues, pour une large part, à la lourdeur inhérente au fonctionnement d'une instance d'évaluation interministérielle réunissant plusieurs dizaines d'acteurs représentant les différents ministères et les acteurs socioprofessionnels concernés par la politique évaluée. Cette lourdeur n'est-elle pas alors, au moins pour partie, le prix à payer pour alimenter de véritables débats contradictoires et dégager un consensus partagé sur l'analyse d'une politique publique, permettant ensuite de plus facilement la réformer ? Le fait est que les dispositifs existant aujourd'hui reposent beaucoup moins sur de telles instances, et que les évaluations réalisées sont beaucoup moins présentes dans le débat public, même dans les cercles d'initiés. Par conséquent, l'objectif de comprendre finement une politique (et qui nécessite souvent un temps plus ou moins long selon la complexité de la politique étudiée) est souvent mis en second plan au profit d'une logique plus utilitariste.

Concernant ensuite les débats au sein de la SFE, les deux dernières thématiques que nous avons mises en avant précédemment renvoient assez directement à des enjeux de méthodes. L'interprétation à faire du principe de pluralité, en premier lieu, et donc le degré de pluralisme des acteurs à associer à une évaluation, fait directement écho à l'exemple de l'évaluation interministérielle. Les tenants d'une vision restrictive de ce

principe mettent eux aussi en avant la lourdeur inhérente à sa mise en œuvre, qui pénalise de leur point de vue l'utilité décisionnelle des évaluations, et donc sa progression dans la sphère publique. Les partisans de l'autre vision insistent pour leur part sur la finalité démocratique de l'évaluation, nécessaire à l'acceptabilité de ses résultats, voire se réclament des principes du développement durable pour défendre une pratique « citoyenne » de l'évaluation.

La question de l'évaluation des impacts, en second lieu, est aussi un débat renvoyant en grande partie à des questions de méthode, l'objectif étant ici explicitement la mesure des impacts escomptés des politiques. En effet, les défenseurs des méthodes permettant d'accéder à une estimation statistiquement robuste de l'effet propre d'une politique publique mobilisent des approches statistiques complexes nécessitant selon les cas soit des bases de données très fournies sur les bénéficiaires, soit une mise en œuvre des dispositifs à évaluer selon des modalités dites expérimentales, soit enfin des modélisations socio-économiques (Duflo et Kremer, 2005). Ces conditions n'étant réunies que dans un nombre limité de situations, le domaine de validité de ces approches reste limité. Les tenants d'une approche évaluative moins centrée sur l'évaluation d'impact mobilisent pour leur part un ensemble plus large d'outils empruntés aux sciences sociales — entretiens semi directifs, enquête par questionnaire, études de cas, groupes de travail,...— pour tenter de réunir un ensemble de présomptions sur les chaînes de causalité entre les moyens mis en œuvre et les effets que l'on peut leur attribuer. Si ces appréciations sont indiscutablement moins robustes sur le plan statistique, elles ont néanmoins trois avantages importants par rapport à l'évaluation quantitative des impacts : (1) leur domaine de validité opérationnelle est beaucoup plus important, (2) elles permettent de répondre à une palette de questions beaucoup plus large que celle de la seule estimation des impacts (cohérence, pertinence,...), et (3) elles peuvent fournir des explications sur les causes de réussite ou d'échec, au delà de la preuve et de la quantification de l'existence ou non d'un impact. La complémentarité entre approches quantitatives et qualitatives apparaît donc essentielle, mais les tentations hégémoniques des premières restent pourtant vives, alimentées par la demande des décideurs préoccupés avant tout de savoir rapidement si un impact de leur politique peut être « chiffré ». Malgré tout, depuis quelques années, ces démarches couplant approches qualitatives et quantitatives sont toutefois développées et étayées par les tenants des approches réalistes, prônant la nécessité, face au développement de modèles statistiques abstraits, des analyses visant la compréhension fine de l'objet étudié et le retour à des formes de pragmatisme scientifique (Pawson et Tilley, 1997 ; Pawson, 2006)

En ce qui concerne enfin l'évaluation des politiques cofinancées par l'Union Européenne, nous avons souligné combien l'importance accordée à l'instruction d'un ensemble d'indicateurs définis en début de programme n'avait cessé d'augmenter. Même si la volonté de disposer d'éléments de comparaison entre Etats membres est compréhensible, voire louable, un certain nombre de voix commencent à s'élever pour souligner les limites de cette approche. Elles sont de deux types :

— sur la méthode elle-même, car la définition d'indicateurs identiques susceptibles de rendre compte de dispositifs déclinés de façon différente dans chaque pays, voire région, est en soi une gageure, et par ailleurs, les protocoles d'instruction étant par nature « déconcentrés », il n'existe aucune garantie sur la comparabilité effective des résultats ;

— sur les effets pervers de cette approche, car cette demande très « centralisée » limite grandement l'appropriation des exercices évaluatifs par les acteurs pilotant les programmes, et assimile l'évaluation à du suivi, de surcroît « imposé », pour rendre des comptes à un financeur.

pour rendre des comptes à un financeur.

On peut tout à fait voir dans cette approche européenne une tendance décrite par ailleurs, dans laquelle l'indicateur dépasse son statut d'outil de mesure d'une réalité objective pour devenir l'instrument d'une nouvelle gouvernance (par le chiffre plus que par le droit) (Zahm, 2011). Le risque est alors de « se focaliser sur la mesure et perdre le sens de la mesure » (Zahm, 2011).

A travers ce deuxième angle d'analyse centré sur les outils et les méthodes, deux tendances se dessinent au sein des trois exemples que nous avons développés. La première est indubitablement l'attraction exercée par le chiffre, tant pour ses vertus de robustesse que de comparabilité. La deuxième, qui en est un peu le pendant, est la mise en doute de la nécessité du pluralisme, toujours accusé de lourdeur, et parfois d'autres défauts comme la place accordée au savoir profane ou le risque d'instrumentation des instances par des lobbys.

6. En conclusion : plus ou moins de pluralisme dans la mesure de la performance des institutions ?

Au terme de cette analyse des tendances observables dans le développement de l'évaluation de politique publique, que peut-on dire à partir de l'exemple de cette pratique sur la dimension plus ou moins pluraliste de la mesure de la performance des institutions ?

Le premier enjeu que ces tendances révèlent est celui du risque d'accorder un poids prépondérant dans l'analyse à certaines « qualités » que l'on peut attendre de l'action publique au détriment des autres. On a souligné en effet les fortes attentes en matière d'efficacité et plus encore d'efficience, qui orientent les doctrines et les méthodes, au risque de délaisser d'autres champs comme ceux de la pertinence ou de l'utilité (la politique répond-elle aux besoins de la société ?).

Cette importance accordée à l'efficience, évidemment liée aux situations de tensions budgétaires qui caractérisent les économies européennes, se fait-elle au détriment des exigences démocratiques ? Les mises en cause du principe de pluralité ou le succès des approches quantitatives pourraient le laisser penser, même si les approches se réclamant en termes de gouvernance du paradigme du développement durable sont également vivaces.

Le deuxième enjeu qui apparaît est celui de la gestion du temps : produire une connaissance partagée et mise en débat sur l'action publique dans ses différentes dimensions a un coût en termes de durée, difficilement compatible avec la rapidité du temps politique. C'est ainsi au nom de la volonté de rendre les évaluations plus « utiles » à court terme pour le décideur que bon nombre des tendances observées sont justifiées. Il n'en reste pas moins que cela revient à accorder une place prépondérante à une

finalité de l'évaluation de politique publique — l'aide à la décision — au détriment des autres : production de connaissance, transparence démocratique, formation des citoyens (Berriet-sollic et al., 2011).

Le risque d'appauvrissement des réflexions est donc réel, si une vision unifiée, largement inspirée des préceptes du « New Policy Management », venait à s'imposer. On en est heureusement encore loin, si l'on en juge par la diversité des pratiques évaluatives encore à l'œuvre au sein des institutions de toute taille qui inscrivent peu à peu l'évaluation dans leur fonctionnement. Mais les tendances que l'on a pu décrire précédemment, d'autant qu'elles émanent d'acteurs institutionnels influents sur l'évolution des pratiques évaluatives, comme l'Etat ou la Commission Européenne, incitent à la vigilance.

Bibliographie

- BARBUT Laurent & BOUNI Christophe — « Évaluation des programmes structurels européens : enseignements à partir de deux exemples contrastés » — Actes des troisièmes journées de l'évaluation — Issy les Moulineaux 2001 — L'harmattan. BERRIET-SOLLIEC, M.; GUERIN, M. VOLLET, D. 2005. L'évaluation des politiques agricoles en France : L'évaluation au milieu du gué ? *Economie Rurale*, N°286-28 : 89-97.
- BERRIET-SOLLIEC, M. ; LABARTHE, P. ; LAURENT, C. ; BAUDRY, J., 2011. Empirical validity of the evaluation of public policies: models of evaluation and quality of evidence. EAAE Seminar "Evidence-based agricultural and rural policy making: methodological and empirical challenges of policy evaluation" ; 2011/02/17-18 ; Ancona (ITA). 2011. 15 p.
- Commission Européenne. 2010, *Conclusions sur le cinquième rapport sur la cohésion*
- Commission Européenne. 2010, *Communication sur la PAC post 2013*
- Commission Européenne, 1999. *Evaluer les programmes socio-économiques*. Luxembourg, Collection MEANS (<http://www.evaled.info>). Conseil scientifique de l'évaluation 1996, *Petit guide de l'évaluation des politiques publiques*, Paris : La Documentation Française
- Conseil Scientifique de l'Evaluation 1995, 1996, 1997, 1998, 1999, *L'évaluation en développement, Rapports annuels sur l'évolution des pratiques d'évaluation des politiques publiques*, Paris : La Documentation française
- DATAR 1998, *Mémento sur l'évaluation des programmes financés par les Fonds structurels européens*
- DAUCE P., 1998. L'évaluation des politiques communautaires de développement régional. Enjeux, méthodes, résultats. L'exemple de l'objectif 5B en Bourgogne, *Revue d'Economie Régionale et Urbaine*, 98(3) : 379-394.
- DUFLO E., KREMER M. (2005). Use of Randomization in the Evaluation of Development Effectiveness." in Pitman G., Feinstein O., Ingram G. (eds.), *Evaluating Development Effectiveness*. New Brunswick (NJ): Transaction Publishers, 205-232.
- GREFFE, X. 1997. *L'évaluation des projets publics*. Paris : Economica.
- MONNIER, E. 1987. *Evaluation de l'action des pouvoirs publics*. Deuxième édition en 1992. Paris : Economica, 199 p.

- NIOCHE, J.P. ; POINSARD, R. 1984. *L'évaluation des politiques publiques*. Paris : Economica.
- PATTON, M. Q. 2008. *Utilization focused evaluation*, California, 4th Edition, Sage.PAWSON R.; TILLEY, N. 1997. *Realistic Evaluation*, London, Sage.
- PAWSON , R. 2006. *Evidence-based policy : a realistic perspective*. London. Sage publications.
- PERRET B. 2001, *L'évaluation des politiques publiques*, Paris : Repères – La Découverte
- ROSSI, P. ; FREEMAN, H.E. 1993. *Evaluation: a systematic Approach*. Sage Publications.
- Shadish, W.R., Cook, T.D. & Leviton, L.C. (1991). *Foundations of Program Evaluation Theories of Practice*. Newbury Park: Sage publications.
- ZAHM F. 2011, *De l'évaluation de la performance globale d'une exploitation agricole à l'évaluation de la politique agroenvironnementale de la PAC : une approche par les indicateurs agro-environnementaux*, Thèse de doctorat soutenue le 6 avril 2011