

Évaluation des politiques publiques et quantification en France : des relations ambiguës et contradictoires entre disciplines

Jean-Claude BARBIER

CES, Université Paris 1

jean-claude.barbier@univ-paris1.fr

Nicolas MATYJASIK

CERAPS, Université Lille 2

matyjasik.nicolas@orange.fr

Le recours à la mesure et à la quantification des politiques et des activités administratives est ancien, et associé parfois à une forme de « reddition des comptes » par les administrateurs. L'évaluation des politiques publiques (EPP) est nettement plus récente et, en France elle peine à trouver une place légitime. Cela tient à la fois à l'histoire des institutions politiques dans ce pays, mais aussi aux relations que l'évaluation entretient avec les trois disciplines (économie, science politique et sociologie) qui la nourrissent. Nous montrons que les rapports sont plus ou moins directs, souvent ambigus et stratégiques.

Strategies for quantifying policies' outcomes and administrative activities have had a long history; they have been associated with the exercise of accountability. On the other hand, programme evaluation in France is much younger. It is not easily accepted in this country. This situation has obviously much to do with the legacy of French political institutions. But it is also largely due to the ambiguous relationships programme evaluation has had with three academic disciplines, economics, sociology and political science.

Mots clés : évaluation • aide à la décision • méthodes quantitatives • politiques publiques

Keywords: programme evaluation • quantitative methods • public policy

Les auteurs du présent texte considèrent que le mariage ne va pas de soi entre les multiples activités : mesurer, quantifier, évaluer. Ce constat ne vient pas de la technicité de chacune de ces

opérations, mais de leur *sociologie*. L'étude proposée ici vaut essentiellement pour le cas français, dont il est connu depuis bien des années qu'il affiche, en matière d'évaluation des politiques publiques (EPP¹), des caractéristiques différentes de bien d'autres pays [Perret, 2001 ; Barbier, 2004, 2007 ; Jacob, 2005 ; Bezes, 2009]. Les rapports difficiles de l'évaluation et de la quantification en France apparaissent dès lors qu'on tente une sociologie de l'EPP, en balayant une histoire relativement récente ; c'est ce qu'on explore dans une première partie, en prenant appui sur la comparaison avec d'autres pays. L'évaluation a des difficultés à se situer parce que l'expertise, quantifiée ou non, est l'apanage de producteurs institutionnalisés et qu'elle ne trouve ni sa place ni ses avocats. La référence à l'évaluation vient de l'étranger – notamment des États-Unis (*program evaluation*) –, et les agents de l'expertise par excellence, ceux de l'Insee, sont restés étrangers à la référence jusqu'aux années récentes. En outre, nous faisons l'hypothèse que la situation française s'explique, aussi en partie, par des rapports malaisés avec les trois disciplines essentiellement concernées par l'EPP, à savoir l'économie², la science politique et la sociologie. L'économie semble n'avoir jamais eu, dans notre pays, de difficulté spécifique avec l'EPP, dans la mesure où elle a tendu à banaliser sa participation à des opérations définies institutionnellement ou non comme de l'évaluation : elle a systématiquement eu des liens avec le calcul économique, et ses liaisons avec l'EPP apparues en France dans les années 1980 peuvent, en quelque sorte, être considérées dans la continuité de ce lien ancien. En revanche, la situation est bien plus complexe du côté de la science politique et de la sociologie qui seront, au niveau de précision envisagé ici, considérées ensemble. En définitive, une analyse disciplinaire fait ressortir des enjeux particuliers à la France, sérieusement marqués par la nature du système politique, mais aussi par les intérêts des communautés disciplinaires et par l'inégalité structurelle des ressources de pouvoir entre elles. Ces enjeux se combinent pour produire un cas français bien différent du cas britannique, pour ne prendre qu'un exemple classique. Non pas qu'il s'agisse en aucune sorte de reprendre le thème rebattu de « l'exception française ». C'est plutôt que mesurer, quantifier, évaluer, ne sont pas des activités *spontanément universelles* en Europe.

1. Histoire de l'évaluation des politiques publiques en France, repères comparatifs

1.1. Repères essentiels

Une *sociologie politique explicitement centrée sur l'évaluation* reste, pour l'essentiel, à faire (voir Monnier, 1992, et Spenlehauer, 1998). Par comparaison avec d'autres pays européens, et, surtout avec le cas américain, l'apparition de l'évaluation en France paraît tardive : à la fin des années 1970 et, surtout, dans les années 1990, elle s'incarne pour la première fois dans des institutions formelles de l'État (central et local) ; parallèlement, les prolégomènes de la naissance d'une identité professionnelle d'« évaluateur », certes fragile, apparaissent ; un marché de l'évaluation se constitue progressivement, dont les produits (ce sont évidemment des services professionnels complexes) se distinguent difficilement d'autres, la recherche, l'expertise, l'étude,

¹ Par commodité, on emploiera « évaluation » tout court pour alléger le texte, alors que l'objet de l'analyse consiste en l'évaluation des politiques et des programmes publics.

² Compte tenu du caractère exploratoire de l'étude présente, on ne fait pas allusion systématiquement à la gestion, discipline qui établit des ponts aussi bien avec la sociologie des organisations qu'avec l'analyse budgétaire, etc.

l'audit, le contrôle, le management, etc. Au cours de la décennie 1990 et dans les premières années 2000, les formes prises par l'évaluation ont évolué. Un premier type d'opérations a été commandité par les financeurs publics visant à objectiver les « impacts » (ou « effets nets ») des actions publiques, qui s'associe souvent à un suivi (*monitoring*) des programmes et à l'analyse de leur mise en œuvre, des systèmes d'acteurs, et des enjeux du management public, en termes d'analyse des politiques publiques [Nioche, 1982]. Plus récemment, en France et ailleurs, des évaluations ont mobilisé la micro-économie néo-classique³.

Quoi qu'il en soit, à la fin des années 1980 et cours des années 1990, un nombre croissant d'acteurs français divers (hauts fonctionnaires, experts, consultants, universitaires, etc.) ont choisi de faire référence à l'évaluation des politiques et des programmes pour qualifier *certaines de leurs activités*. C'était nouveau. Progressivement, bien que de manière très inégale, l'EPP a commencé d'être repérable parmi les activités spécialisées liées à la mise en œuvre et à la conception des politiques. D'où un processus d'*institutionnalisation* [Monnier, 1992 ; Jacob, 2005], visible, notamment, dans le fait que des administrations de l'État (central et local) l'ont eue en charge. L'EPP s'est progressivement installée dans le paysage politique français dans les années 1990, formellement, sous l'incitation de la circulaire de 1989 du Premier ministre Rocard⁴. Cette période connaît un ensemble assez disparate d'initiatives, dont certaines étatiques, menées par une pluralité d'acteurs. L'évaluation connaît un certain succès parmi d'autres cercles qui s'occupent de gestion publique [Bezès, 2009, p. 267-273]. Une controverse latente et parfois explicite opposa les tenants d'une évaluation « démocratique » à une « évaluation managériale ». Une évaluation dite « pluraliste » prit son essor. L'activité d'évaluation connaît une visibilité importante dans l'évaluation de la loi expérimentale du revenu minimum d'insertion (RMI, 1988-1992) et dans la discussion sur le rapport commandé à P. Viveret (1989). Cette « tradition » d'évaluation est éclectique, du point de vue des disciplines, des méthodes d'évaluation, des commanditaires, des pratiques professionnelles et des acteurs.

À la fin des années 1980 et dans les années immédiates qui ont suivi, il a pu sembler aux acteurs qu'un pas décisif avait été franchi dans l'institutionnalisation de l'évaluation en France, avec la création, en 1990, du Conseil scientifique de l'évaluation (CSE). Dans son activité de surveillance et de suivi des évaluations (baptisées « interministérielles ») commanditées par les grands ministères, le CSE contribua à produire les principes d'une « doctrine ». Bien qu'il n'ait jamais réuni les ressources adéquates pour asseoir une influence prépondérante, ce dernier n'en constitua pas moins un référent majeur. L'évaluation constitue l'une des marques de la « configuration de réforme » de la période [Bezès, 2009]. Ensuite, le CSE est tombé en désuétude, puis a été remplacé, en 1998, par une instance différente, le Conseil national de l'évaluation (CNE). Pendant que, de son côté et sans coordination de fait, le marché se développait, en particulier à travers les commandes croissantes des collectivités territoriales, le CNE tomba finalement lui aussi en sommeil et ne fut jamais renouvelé. À partir des années 2000, une autre dynamique de la « modernisation » de l'État lui succéda. Avant d'y

³ Voir sur ce point Nioche (1982) qui en note les limites. Pour les États-Unis, récemment : [C.H. Weiss et al., 2008].

⁴ Circulaire du 23 février 1989.

revenir, il faut entrer plus précisément dans les origines de la *démarche évaluative*, qui se présente comme une activité spécifique, certes pas comme démarche de quantification *mainstream*, ni comme expertise classique, mais comme activité liée, de façon plus ou moins explicite selon les conceptions de ses acteurs, au management public, entendu comme l'ensemble des pratiques d'organisation et de gestion de l'État.

1.1.1. De l'inspiration intellectuelle américaine à la doctrine du Conseil scientifique de l'évaluation

Pour schématiser, l'activité baptisée EPP en français dans les années 1990 puise son inspiration dans la pratique américaine de *program evaluation*, particulièrement florissante au moment des programmes sociaux sous les présidences Kennedy et, surtout, Johnson. Elle plonge ses racines intellectuelles aux États-Unis et ailleurs, bien plus anciennes, dans les pratiques de *social engineering* des débuts du *xx^e* siècle, en particulier dans les expérimentations pratiquées dans le domaine de l'éducation. En France, une longue tradition d'ingénieurs économistes-statisticiens peut également être rattachée à la généalogie de l'évaluation telle qu'elle se présente au début des années 1970, sous un avatar, dit « rationalisation des choix budgétaires » (RCB). Globalement on peut considérer que l'expérience échoua, puisque la RCB fut abandonnée finalement en 1985.

Une nouvelle étape du processus d'acclimatation de l'évaluation en France fut marquée par la publication, en 1986, d'un rapport du Commissariat général du Plan [Deleau *et al.*, 1986]. Très globalement, ce texte s'inscrivait dans la réflexion internationale de l'époque en termes de « crise de l'État-providence », ou des « limites du *welfare state* » [Duran *et al.* 1995, p. 47], et, plus généralement, de ce qui allait nourrir le mouvement international dit du nouveau management public (*New public management*) ; son orientation centrale défendait la nécessité d'évaluer les politiques françaises du point de vue de l'efficacité ou de l'efficience, vis-à-vis des coûts (coût-bénéfices, coût-avantages). Cela lui fut reproché par d'autres partisans de l'évaluation en France. L'impulsion donnée par le gouvernement de gauche, sous le Premier ministre Rocard, en 1989, fut, au moins en partie, différente, sensible qu'il était à l'argumentation du rapport Viveret, commandité par lui. L'évaluation des politiques publiques y était vue comme l'un des éléments d'une « modernisation de l'État ». Pour autant, des activités qualifiées ici ou là d'évaluation n'en continuèrent pas moins à se développer, à la frontière d'autres activités plus classiques (inspection, étude). Dès cette époque, on voit s'affronter en France des milieux qui portent des conceptions politiques divergentes. Alors que les caractéristiques concrètes de l'activité qui se nomme *évaluation* ne séparent pas, sans chevauchement, cette dernière d'autres activités liées au management public, la description sociologique, en France ou à l'étranger, ne peut cependant ignorer les modalités par lesquelles les milieux de l'évaluation se sont construits, les intentions des acteurs, de même que les programmes politiques qu'ils portent, au-delà des affrontements classiques entre groupes d'intérêts. Cependant, le travail d'objectivation sociologique des projets présents dans les « milieux de l'évaluation » reste à faire.

S'il faut en croire sa composition initiale, le CSE fut, dans un premier temps, considéré comme un enjeu relativement important par la haute fonction publique : de hauts

fonctionnaires des grands corps d'inspection, de l'Insee et de la Cour des comptes côtoyaient, en nombre à peu près équivalent, des universitaires et experts, et un seul représentant du secteur privé au départ. Il n'est certes pas indifférent que le Conseil ait été présidé par J. Leca, professeur internationalement renommé de science politique. En fait, après des débuts actifs, la procédure de sélection de projets s'embourba et il n'y eut plus, à partir de 1995, de nouveaux projets retenus ; dans le même temps, les membres du CSE qui le quittaient ne furent pas remplacés par les autorités compétentes. Tout se passait comme si l'enjeu lié à l'évaluation, dans la haute fonction publique s'était radicalement amoindri [Bezes, 2009]. Au total, pendant la période 1990-1996, le dispositif CSE n'aura choisi et suivi qu'une vingtaine d'évaluations dont la plupart étaient certes de grande envergure et pleines de leçons [Barbier, 2004], mais qui ne représentaient qu'une petite partie de l'activité existant sous l'étiquette « évaluation » en France à l'époque. En effet, dans le même temps, le rôle des conseils régionaux comme commanditaires d'évaluations se renforçait, sous l'impulsion du cofinancement, par la Commission européenne, des actions relevant des fonds structurels (notamment le FSE et le Feder). Une autre source d'opérations évaluatives fut fournie par la création, en 1993, de l'évaluation obligatoire des « contrats de plan État-régions⁵ ». En dépit de leur importance à de nombreux égards, on a pu considérer que, du point de vue d'ensemble de la position de l'EPP en France, ces opérations restaient modestes [Spenlehauer et Warin, 2000].

1.1.2. L'évaluation et la réforme de l'État

À partir de 1998, le CNE a fait conduire une nouvelle série d'évaluations, dont les résultats ne furent pas tous publiés [Barbier, 2004]. Après 2002, le gouvernement n'ayant pas renouvelé le Conseil, l'État n'a plus disposé d'instance centrale. Au lieu de cela, il a favorisé des démarches qu'on peut considérer comme apparentées à l'EPP. Celles-ci ont en quelque sorte « remplacé » l'évaluation telle que précédemment pratiquée et, surtout, telle que promue par feu le CSE. Ces initiatives s'inscrivirent bien évidemment dans un mouvement plus long de réforme de l'État, dont il n'est pas possible de rendre compte ici, faute de place [Bezes, 2009]. Dans la période la plus récente, tout d'abord, les gouvernements ont conduit, dans les différents départements ministériels, des « stratégies ministérielles de réforme », à partir de 2003. Ensuite, la procédure budgétaire a été entièrement refondue sous l'impulsion de l'adoption (en août 2001) puis de la mise en œuvre systématique depuis, de la loi organique relative aux lois de finances (Lof) : cette réforme d'importance majeure 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. Ses premiers initiateurs laissaient entendre que la Lof allait satisfaire le besoin d'une EPP, ce qui n'a pas été le cas, comme le prouve, en particulier, la persistance structurelle du financement d'opérations d'évaluations sectorielles dans de nombreux départements ministériels. Enfin, de manière plus conjoncturelle, le gouvernement nommé par le nouveau président de la République, en 2007, comprenait un secrétaire d'État dont les compétences comportaient explicitement « l'évaluation des politiques publiques ». Le président français a aussi lancé

⁵ Circulaire (9 décembre 1993) « relative à la mise en œuvre de la démarche d'évaluation dans les procédures contractuelles (contrats de plan – contrats de ville) ».

une opération ponctuelle qu'il a dite inspirée de la pratique canadienne, et qu'il a nommée « révision générale des politiques publiques – RGPP ». Alors que cette dernière, à l'origine, a largement puisé dans le répertoire habituel de l'évaluation, son articulation avec les pratiques plus anciennes d'EPP n'était pas évidente. Au moment où la première phase de la RGPP est parvenue à son terme (2009), il est clair que, pas plus que la Lof, ces deux procédures n'ont effectivement pénétré sur les terres de l'évaluation des politiques publiques, dans son sens le plus consensuel, accepté dans la communauté internationale composite des auditeurs et évaluateurs [Furubo *et al.*, 2002]. En revanche, les plus récents changements en cours sont liés à la modification de la Constitution, en été 2008 et à l'interprétation qu'en fait, notamment, l'un des acteurs essentiels, à savoir la Cour des comptes. Un an après, il est évidemment trop tôt pour en imaginer l'impact sur la pratique de l'EPP en France.

Cette brève présentation illustre la présence malaisément reconnue des activités dites d'évaluation en France. Dans d'autres pays où les controverses ne sont évidemment pas absentes, la situation est cependant souvent différente. C'est surtout le cas dans ceux où des pratiques et références professionnelles sont largement reconnues, qui donnent crédibilité à une *existence distincte* de l'activité d'EPP, appuyée sur des communautés professionnelles relativement identifiables.

1.2. Des traits spécifiques de la situation française en comparaison internationale ?

Sans envisager ici une analyse approfondie de l'idée et des pratiques d'évaluation en France, qu'il faut replacer dans une genèse longue au sein des rivalités pour la représentation du monde parmi les acteurs des politiques publiques [Jobert, 1992], on peut souligner quelques traits qui contribuent à expliquer cette situation.

1.2.1. Caractères et culture politiques du système français d'expertise

Au titre des caractéristiques du système politique français, il faut d'abord signaler la faiblesse du Parlement face à l'exécutif (au moins depuis 1958), ce qui explique son investissement marginal dans les questions de l'évaluation, malgré l'existence d'organes spécialisés⁶. On peut aussi ajouter la pratique du cumul des mandats pour les hommes et femmes politiques, qui aboutit à ce que l'investissement dans les arènes autres que nationales soit toujours partiel et relatif, les hommes et femmes politiques d'importance étant tous visibles dans l'arène nationale : l'opposition, par exemple, entre élus régionaux et État central est toujours relative, ce qui diminue la demande d'évaluations contradictoires entre niveaux de gouvernement légitimes. Cet état de choses se trouve congruent avec l'organisation et l'histoire de l'administration française, qui, *structurellement*, présente une tendance à exclure l'évaluation au bénéfice du contrôle et de l'expertise de l'État *central*, lesquels aiment à se légitimer, dans une tradition jacobine, comme les porteurs exclusifs de l'intérêt général. Comme l'a mon-

⁶ Avant la réforme constitutionnelle de l'été 2008, plusieurs « offices parlementaires d'évaluation » étaient spécialisés, ainsi qu'une mission d'évaluation et de contrôle (MEC). Ce système est bouleversé par la mise en place du nouveau rôle du Parlement dans l'EPP.

tré un rapport rédigé par des sénateurs [Bourdin, 2001]⁷, qui comparait les situations américaine et française, on est en présence en France d'un quasi-monopole des systèmes statistiques, des études et de l'expertise : s'il ne s'agit d'un quasi-monopole de production que dans le cas des services de l'Insee, l'expertise ou l'évaluation ont beau être confiées à des consultants ou des universitaires⁸, elles ne se font qu'exceptionnellement dans un cadre de commande indépendante [Jobert, 1994]. Pour les raisons indiquées précédemment, le poids des régions et autres collectivités territoriales n'est pas en mesure de contrebalancer ce quasi-monopole. La thématique de « l'intérêt général » permet d'illustrer la situation française par comparaison avec les autres pays développés. L'attitude vis-à-vis de cet « intérêt général », sa représentation, sa rhétorique, distinguent les sociétés en Europe en tant qu'un des éléments de cultures politiques restées nationales malgré l'europanisation des élites. Par exemple au Royaume-Uni et dans le monde anglophone ou scandinave, le système politique est évalué, au moins au niveau de la rhétorique, au regard de son « *accountability* », c'est-à-dire de la façon dont il se soumet à la reddition des comptes de son action au nom de l'intérêt public. Dans la tradition républicaine pluriséculaire, les agents de l'État doivent certes rendre des comptes aux citoyens, comme le précise l'article 15 de la Déclaration des droits de l'homme et du citoyen de 1789. Mais le système français se caractérise par le rôle éminent joué par les « grands corps » de l'État dans les tâches d'inspection, d'expertise et de contrôle [Bezes, 2009]. Une telle organisation fondée sur l'expertise étatique (comme celle des inspecteurs des finances ou des administrateurs de l'Insee) est étrangère à la construction d'une connaissance partagée avec la société civile, puisque les agents du contrôle y sont censés représenter des incarnations de l'intérêt général [Jobert 1992]. Pour ce qui concerne la reconnaissance de l'EPP, aux côtés des activités « installées » (contrôle, inspection, etc.), les grands corps détiennent inévitablement un pouvoir de contrôle (de *gate-keeper*) assis sur des ressources institutionnelles et sociales ancrées dans le système politico-administratif. Pour que l'évaluation se développe en tant que champ particulier d'une profession, concurrente et complémentaire à la fois du contrôle ou de l'inspection, sur la base de standards indépendants, il faudrait logiquement que ces acteurs traditionnels y trouvent un intérêt, ou que la dynamique d'un changement soit impulsée « de l'extérieur », par exemple par la concurrence du secteur privé. L'expérience française montre que ces conditions n'ont pas encore été réunies, alors même que dans la moitié des années 1990, une partie de la haute administration s'est résolument tournée vers le *New public management* [Bezes, 2009] et le référentiel « néo-libéral gestionnaire » [Jobert, 1994].

1.2.2. La place centrale des statisticiens et l'absence d'un métier d'évaluateur

En conséquence, du côté des corps d'inspection et de statisticiens, l'idée d'une spécificité de l'évaluation reste peu reconnue, relativisée, parfois ignorée ou même combattue. L'ignorance de l'évaluation est, encore aujourd'hui « incorporée » dans

⁷ Ceux-ci ont noté qu'il n'existait en France à l'époque que quatre petites organisations indépendantes productrices d'expertise (dont l'OFCE est la plus connue) [Bourdin, 2001, p. 18-25]. En Allemagne par exemple, une grande variété de fondations contribue à la production d'une expertise pluraliste.

⁸ On ne souligne pas assez que les universitaires français sont aussi des « hauts » fonctionnaires.

l'éthos traditionnel des agents de l'Insee [Barbier, 2008]. Ces faits doivent être compris comme des manifestations d'une certaine culture politique française. On peut aussi ajouter que les universitaires n'ont pas plus d'autonomie dans ce jeu stratégique. On peut ainsi considérer, dans le cas français, que le rôle de l'Insee, institution par excellence liée à la quantification en France, a joué de façon défavorable à l'implantation de l'évaluation dans ce pays. En outre, la structuration d'un milieu professionnel spécifique n'a pas non plus (pas encore ?) résulté en France de l'influence du secteur privé qui occupe une place dominante parmi les producteurs d'évaluation. De ce côté, l'offre de conseil, sérieusement segmentée, n'est pas centrée sur l'EPP, un produit qui reste marginal pour elle. Les grands cabinets internationaux de conseil en management, contrôle de gestion, comptabilité et audit ont certes accru considérablement leur présence sur les marchés publics au niveau central et territorial. Mais ce phénomène a concerné surtout le management public, l'audit, l'organisation, à la faveur, comme en d'autres pays, de l'influence du *New public management*. En matière d'évaluation, ces grands cabinets, trouvant le marché public incertain et d'une rentabilité inférieure, y ont participé inégalement et diversement selon la conjoncture du marché privé. De l'autre côté, les offreurs de produits d'évaluation en France se divisent en cabinets moyens et petits qui réalisent des études à visée évaluative dans le prolongement de leurs niches construites à partir d'activités d'études sectorielles, de contrôle de gestion ou de conseil [Matyjasik et Méasson, 2006]. Certains acteurs professionnels de l'évaluation en France ont pensé que l'eupéanisation allait entraîner une assimilation des pratiques de l'évaluation en France [Toulemonde, 2000]. Mais, jusqu'à présent, le marché de l'évaluation, que l'Union européenne impulse il est vrai de façon décisive, a pu se développer sans que, pour autant, des coalitions d'acteurs suffisantes se forment pour introduire, au sein du système politico-administratif relativement hostile qu'on a décrit, des conceptions professionnelles équivalentes à celles qui ont cours dans de nombreux autres pays. Notamment, l'offre de formation supérieure peine à se développer en 2009, ainsi que les références distinctives de la construction d'un « métier ». On notera cependant qu'en 1999 le « monde » français de l'évaluation s'est enrichi d'un nouvel acteur, la Société française de l'évaluation (SFE), dont l'article 3 des statuts fixe la mission de « contribuer au développement de l'évaluation et de promouvoir son utilisation dans les organisations publiques et privées. Elle se fixe notamment pour objectifs de faire progresser les techniques et méthodes et de favoriser le respect de règles éthiques et procédurales propres à garantir la qualité des évaluations ainsi qu'un usage approprié de leurs résultats »⁹.

1.2.3. Controverses et évaluation dite « pluraliste »

Une autre particularité de la situation française réside dans les controverses à propos de l'EPP. P.A. Sabatier (1998, p. 143) souligne que le cœur de la question relève de la possibilité (ou de l'impossibilité) de la construction, dans une communauté d'évaluateurs, de standards professionnels, autour desquels les différents groupes se trouvent contraints de contester les conceptions de leurs adversaires, et de se mettre d'accord

⁹ La SFE a, en 2003, adopté une « Charte de l'évaluation des politiques publiques et des programmes publics ». Les principes sont ceux de : pluralité (prise en compte des différents points de vue) ; distanciation (autonomie et impartialité de l'évaluation) ; compétence (des participants) ; respect des personnes (protection de leur intégrité et sécurité) ; transparence (justification des conclusions et diffusion publique) ; opportunité (justification de l'utilité) ; responsabilité (division des rôles et responsabilité conjointe des participants).

sur des conceptions communes de l'évaluation, scientifiques et professionnelles¹⁰. Or, aujourd'hui, le champ des activités et acteurs se réclamant de l'EPP est, malgré plus d'une décennie d'histoire, marqué en France par une grande variété de conceptions, de méthodes et de points de référence. Cet éclectisme fait que la question des frontières entre l'EPP et d'autres activités, présente dès les années 1990 a été sans cesse remise sur le métier.

Toutefois, l'éclectisme et la variété ne s'expliquent pas uniquement en termes de champs professionnels plus ou moins différenciés. La controverse politique fut initialement définie par certains de ses protagonistes comme l'opposition entre une conception « démocratique » de l'évaluation, s'opposant à une conception dite « managériale ». Une tension, en partie fonctionnellement analogue à celle qu'on a pu connaître aux États-Unis avec la présidence Reagan, explique la querelle. Cette dernière en dit long sur une partie traditionnelle de la culture politique française qui tend à penser le « management » ou la gestion comme, par principe, étrangers à une mission supérieure de la politique, laquelle ne saurait s'embarasser de questions prosaïques, parce qu'elle tire sa valeur d'une orientation à laquelle des hommes et femmes politiques se réfèrent, dans le vocabulaire français, comme « volontariste ». Dans les termes de la polémique politique, l'évaluation « managériale » d'alors était de droite, alors qu'elle s'appelait « aide à la décision » à l'autre bord, et que l'évaluation « démocratique » était inévitablement de gauche pour ses partisans, alors que ses adversaires tenaient que la véritable évaluation démocratique serait celle qui informe l'électeur sur l'efficacité et les coûts des politiques, avec reddition de comptes.

L'approche de l'évaluation dite « pluraliste¹¹ » est restée également l'objet de controverses. Sa définition a été élaborée progressivement et plusieurs auteurs et institutions y ont participé. E. Monnier y consacre de longs développements dans son ouvrage pionnier (1987 ; 1992) : sa définition tisse ensemble le *constat analytique* selon lequel le déploiement d'une politique publique est toujours le résultat d'interactions multiples entre des acteurs sociaux, avec l'*appréciation normative* selon laquelle la meilleure évaluation est pluraliste, car, négociée et associant les partenaires pertinents d'une politique publique, celle-ci permettrait une meilleure efficacité et légitimité des opérations d'évaluation (*ibid.*, p. 123-124). De manière moins explicite, c'est également le choix privilégié le plus souvent par l'ex-CSE qui a vu, dans l'installation d'une instance pluraliste, représentant les divers points de vue qu'il qualifiait de « légitimes » à propos d'une politique, la meilleure stratégie cognitive (CSE, 1992, p. 14). Une telle instance est porteuse de légitimation pour l'évaluation, ont remarqué des sociologues, d'un point de vue à la fois politique et scientifique [Lascoumes et Setbon, 1996, p. 10-11]. Après de longs débats en son sein, La SFE a également repris cette préférence pour l'évaluation pluraliste, jusqu'à faire de la « pluralité » le premier principe de sa charte, adoptée en 2003.

En définitive, on peut noter l'influence durable, quoique minoritaire, de ce qu'on peut appeler la « doctrine du CSE » dans la situation française. Le rôle de J. Leca dans sa

¹⁰ Voir aussi sur ce point Leca in CSE, 1992, p. 13.

¹¹ Présentée à des lecteurs internationaux, elle a été qualifiée d'évaluation « à la française » par Duran et al., (1995) ; voir aussi : [Toulemonde 2000, p. 356].

création a été vraisemblablement déterminant, et il n'est pas inutile de rappeler le point principal qu'il a systématiquement argumenté. Celui-ci tient dans un postulat. Pour lui, parce que l'information est un instrument à caractère stratégique, « l'évaluation doit être conçue comme une zone d'autonomie par rapport aux acteurs (y compris l'État) auxquels elle apporte de l'information. Sinon, elle devient, dans son déroulement même (et non pas seulement dans ses usages ultérieurs), un élément même de la manipulation de l'information que les stratèges *ne peuvent que* pratiquer » [Leca, 1993, p. 189-190]. L'essentiel semble donc reposer, dans la conception de J. Leca, sur la possibilité de la construction de cette zone d'autonomie, à l'écart, nous dit-il, du système « politico-administratif » [Leca, 1997]. Or, cette possibilité n'est pas donnée et, c'est le moins qu'on puisse dire, elle est difficile à construire par des acteurs qui ont bien d'autres intérêts à défendre que cette *orientation civique* de renonciation, ne fût-ce que pour un temps, à la « polémique » et à la manipulation de l'information. Il nous faut maintenant approfondir l'analyse en interrogeant les rapports particuliers qu'entretiennent les disciplines avec l'évaluation. Ici, l'économie, d'un côté, la sociologie et la science politique, de l'autre, ne sont pas, comme on l'a déjà indiqué précédemment, dans une position analogue. L'économie est naturellement liée à la quantification. La sociologie et la science politique rencontrent des questions épistémologiques communes.

2. L'évaluation et ses disciplines

Plusieurs recherches ont mis en exergue la formation des liens entre les sciences sociales et l'administration en France dans la seconde moitié du ^{xx}^e siècle, en qualifiant de « sciences de gouvernement » [Ihl, Kaluszynski et Pollet, 2003] les savoirs universitaires « investis » par les autorités publiques dans le dessein de donner davantage de scientificité et par conséquent d'autorité à leurs pratiques. Dans la tradition de « l'esprit » des ingénieurs économistes [Vatin, 2008], l'économie s'est très rapidement saisie de l'évaluation pour y apporter les méthodologies forgées par le calcul économique et offrir ses services aux pouvoirs publics. Au contraire, du côté de la science politique et de la sociologie, les relations semblent surtout tenir à l'implication de certains entrepreneurs moraux de l'évaluation qui ont défriché ce champ avant de s'en détourner. En tous cas, à la différence de la situation américaine par exemple, on admettra, dans le cas français, qu'il est impossible d'appréhender l'évaluation comme une « discipline », mais qu'on doit plutôt la considérer en tant que champ d'intervention où quelques acteurs universitaires sont positionnés (parfois explicitement comme évaluateurs, mais cela ne nous paraît pas être le cas majoritaire quelle que soit leur discipline), tout en laissant un espace large pour que d'autres praticiens édifient l'espace professionnel émergeant évoqué en première partie. Un clivage se dessine entre « travailler à l'intérieur de » et « travailler sur » l'évaluation [Barbier, 2008, p. 229-235]. L'activité d'évaluation est, dans le premier cas, identifiée comme une pratique de recherche appliquée, tandis que dans l'autre cas, elle est plutôt analysée comme un traceur permettant de comprendre les mutations de l'action publique contemporaine [Lascoumes, Le Galès, 2004] ; voire, tout simplement un terrain de recherche qui suppose une interprétation ultérieure par le sociologue ou le politiste qui admettent, un moment, de participer à l'opération, sachant qu'il ne s'agit pas, pour eux, d'un travail valorisable sur le plan académique.

2.1. L'économie : évaluer pour mieux décider

Le positionnement de l'économie vis-à-vis de l'évaluation démontre un continuum de pratiques qui vont du calcul économique public aux travaux plus récents de chercheurs qui se réclament de l'économie publique. Les relations paraissent avoir été nouées en ligne directe même si l'enseigne « évaluation des politiques publiques » est minorée par les économistes au profit d'appellations qui font surtout référence à l'aide à la décision publique [Greffe, 1997].

2.1.1. Le calcul économique public et la Rationalisation des choix budgétaires (RCB)

Historiquement la quantification des impacts des décisions publiques et des activités économiques est au cœur du travail de la théorie économique et des tenants de l'économie publique. Dès l'origine, les ingénieurs économistes de l'École des Ponts et Chaussées ont été présents pour construire un corpus instrumental du calcul économique public. Ainsi, dans la filiation de J. Dupuit, le calcul économique public peut être défini comme l'ensemble des méthodes permettant d'évaluer l'opportunité collective d'un projet public en mesurant les avantages et les coûts. Dans ce prolongement, si on s'attache à une histoire récente, des théoriciens de l'économie du bien-être tels que M. Allais ou des ingénieurs français sous l'impulsion de M. Boiteux ont construit les méthodes de l'analyse coûts avantages (ACA). Fort de l'élaboration de cette boîte à outils, le calcul économique continua de se développer avec l'introduction de la RCB dans les services de l'État (comme nous l'avons montré dans la section précédente). L'épisode de la RCB, largement inspirée et nourrie de l'expérience américaine du PPBS américain (*Program Planning Budgeting System*) fera pourtant très rapidement l'objet de querelles [Perret, 2001].

Pourtant, même si la RCB échoua dans son entreprise, elle permit d'introduire une certaine forme d'évaluation au sein de l'État et contribua à une première collaboration systématique entre les tenants du calcul économique public et les acteurs décisionnels centraux¹². Des connexions sont alors créées et vont connaître un regain d'intérêt dans les années 1990 au sein du Commissariat général au Plan (CGP) qui considère le calcul économique comme ce qu'il y a de mieux [Boiteux, 1994, p. 113]. Des commissions de travail et des collaborations plus stables sont alors mises en place en son sein et vont participer à l'ancrage cognitif et normatif de l'évaluation dans les institutions administratives.

2.1.2. Un terrain d'application pour de nombreuses branches de l'économie (et de la gestion)

D'une manière générale, l'évaluation produit un cadre d'analyse sous la forme de questions posées par des commanditaires et des décideurs publics : leurs questions sont modélisables en termes économiques quantitatifs, et se trouvent ainsi constituer des études d'économie appliquée, voire de gestion appliquée, dans de très nombreux domaines. Les travaux des tenants de l'économie publique accordent une attention

¹² La RCB a permis de sensibiliser l'administration et les décideurs politiques à une forme nouvelle de savoir ; elle a facilité l'accès ultérieur de l'évaluation aux cercles décisionnels.

particulière à l'éclaircissement des techniques d'évaluation que sont notamment le taux de rendement socioéconomique et les formules d'évaluation multicritères [Baslé et Malgrange, 2006]. Cet intérêt de l'économie publique pour l'évaluation se repère dans la pratique d'universitaires qui interviennent sur les terres de l'aide à la décision en conduisant des études évaluatives, forme d'économie publique appliquée, ce qui va parfois jusqu'au développement d'outils standardisés¹³. Quant à elle, la micro-économétrie, du marché du travail par exemple, entreprend, depuis plusieurs années, des études appliquées qui sont directement nécessaires au développement de son objet de recherche, à savoir l'évaluation de l'efficacité des politiques publiques. Cette dernière peut se faire en particulier en mobilisant des études d'impact fondées sur des techniques de comparaison d'échantillons de personnes « traitées » par un dispositif public et de personnes « non traitées » [Behaghel *et al.*, 2008], mais elle peut aussi modéliser les effets anticipés par la mise en œuvre d'une mesure de droit du travail ou de la diminution de la durée du travail. Pour les partisans de l'existence d'une activité professionnalisée spécifique, ces derniers économistes interviennent donc en tant qu'*évaluateurs* depuis de nombreuses années, sans, la plupart du temps, se présenter explicitement comme tels : ils agissent alors en tant qu'experts ou économistes plus ou moins renommés. L'implication des économistes dans des opérations d'évaluation a pris récemment une nouvelle dimension dans le paysage scientifique français au sein de l'École d'économie de Paris (EEP). Cette dernière redécouvre et promeut aujourd'hui l'évaluation expérimentale aléatoire¹⁴ née dans le champ de la psychologie et de la sociologie aux États-Unis à la fin des années 1960.

2.2. Science politique et sociologie aux marges de l'évaluation

Alors que la science politique aux États-Unis s'est très tôt confrontée à l'évaluation, le cas français montre qu'il n'y a pas eu d'engouement des chercheurs spécialisés en analyse des politiques publiques. La contribution, pourtant importante, de ces deux disciplines tient seulement à quelques individus symboliquement situés.

2.2.1. La *policy analysis*, fer de lance de l'évaluation aux États-Unis

L'analyse rationnelle des politiques publiques mise en avant par l'aventure du PPBS doit être comprise à l'aune de la structuration d'une *policy analysis*, constituée à des fins d'aide à la décision et de conseil aux gouvernants, qui prend son essor aux États-Unis au milieu du siècle dernier. Ainsi, en 1958, C.E. Lindblom développe le projet d'une *policy analysis* afin de fournir des « recettes du bon gouvernement » à des acteurs publics dont l'action est fortement sous contraintes économiques et tension démocratique. Ce courant a pour objectif d'expliquer les échecs des politiques publiques pour en permettre une amélioration. Plus exactement, une attention particulière est portée à cette époque à la rationalisation des politiques de l'« État-Providence »¹⁵.

¹³ On peut citer l'exemple de l'outil *Sevalpro* de « pilotage optimisé de l'action publique ».

¹⁴ En coopérant parfois avec des acteurs historiques de cette époque, appartenant à la *Manpower Demonstration Research Corporation*, fondée en 1974, comme J. Gueron.

¹⁵ La référence aux limites de l'action efficace de ce qu'ils appellent « l'État-providence » est également constante chez les économistes publics. Voir, par exemple : [Greffé, 1997 ; Mény et Thoenig, 1989].

Les imbrications entre évaluation et analyse des politiques sont donc historiquement nombreuses et directes. E. Monnier explique ainsi que l'évaluation s'appuie à l'évidence sur l'analyse des politiques publiques dans la mesure où elle en permet l'intelligibilité [Monnier, 1992]. Il entend par là que ce courant disciplinaire a permis d'introduire des concepts (mise en œuvre¹⁶, théorie d'action,...) éclairants pour l'évaluation. De son côté, V. Spenlehauer invite également à garder à l'esprit que « les liens de consubstantialité » entre les deux approches doivent toujours rester en « toile de fond » de l'étude des pratiques d'évaluation [Spenlehauer, 2006]. Néanmoins, si cette parenté est à relever aux États-Unis, l'analyse des politiques publiques et l'évaluation dans le cas français arpentent des itinéraires éloignés bien que les importateurs de ce courant de recherche aient tenté de lui imprimer un rythme proche [Thoenig, 1976]. Ainsi, l'analyse des politiques publiques qui connaît en France un « âge d'or » [Hassenteufel et Smith, 2004] ne se place pas dans une position d'aide à la décision à destination des pouvoirs publics. Elle s'ancre plutôt dans une conception sociologique de l'action publique qui met en lumière la boîte noire de l'État [Muller, 2000]. D'où des liaisons plus ou moins indirectes et ambiguës entre évaluation et science politique.

2.2.2. En France : le rôle de quelques politistes, entrepreneurs moraux de l'évaluation

Malgré un faible engouement chez les politistes (notamment des partisans de l'analyse des politiques publiques), l'arrivée de l'évaluation dans le champ politico-administratif français a suscité l'intérêt de quelques universitaires qui ont su y trouver une place d'observation participante pour appréhender de l'intérieur les mécanismes de production de l'action publique [Fontaine et Warin, 2000]. Au-delà de cette position privilégiée, les collaborations mises en place ont contribué à légitimer l'EPP dans l'espace institutionnel. À cet égard, les renommées scientifiques de J. Leca (cf. première partie), puis, de façon bien différente, de J.-C. Thoenig [successivement présidents du Conseil scientifique de l'évaluation (CSE)] ont pu jouer, semble-t-il, comme une forme de caution scientifique à une évaluation dont l'institutionnalisation avait précédé la pratique. Implication telle que ces derniers se sont placés en « militants » de l'évaluation en signant une tribune engagée dans *Le Monde* : « L'évaluation ne doit pas être réservée aux seuls experts et décideurs. Elle sera diffusée aux citoyens pour que ceux-ci s'impliquent dans les choix publics (...) L'orientation générale dont a besoin l'évaluation en France aujourd'hui est l'affirmation d'une volonté politique d'en étendre l'usage public. Car l'État ne s'évalue pas naturellement » [Leca et Thoenig, 1997]. Ce rôle structurant dans le développement de l'évaluation fut, par la suite, clairement affiché par J.-C. Thoenig (2006) quand il évoqua son implication au sein du CSE.

La contribution de ces politistes n'est donc pas négligeable et leurs écrits ont bien souvent nourri théoriquement les praticiens de l'évaluation en France (on pense notamment ici à la notion de zone d'autonomie évoquée plus haut). Leur position centrale au sein d'une institution qui a construit un important corpus méthodologique a ainsi joué dans le design d'une approche évaluative dont « l'évaluation pluraliste »

¹⁶ On pense au très important ouvrage intitulé simplement *Implementation*, de J.L. Pressman et A. Wildawski, paru en première édition en 1973.

constitue une caractéristique qui paraît empiriquement dominante [Lascombes et Setbon, 1996], bien qu'elle soit contestée. Par ailleurs, au-delà des cercles centraux, d'autres politistes ont participé au mouvement de reconnaissance de l'évaluation en œuvrant dans des instances locales qui ont permis de stabiliser l'évaluation dans l'espace régional [Spenehauer et Warin, 2000]. À titre d'exemple, on note le cas de J. Fontaine (en Bretagne) ou encore P. Warin (en Rhône-Alpes) : par leur implication locale, ils ont permis, dans une certaine mesure, d'asseoir scientifiquement l'évaluation et de la faire gagner en audience auprès des pouvoirs publics locaux. Les participations évoquées, quoiqu'importantes pour l'EPP en France, se révèlent pourtant une exception plutôt liée à des acteurs isolés, bien qu'influents. Les positions symboliques occupées par quelques politistes ne doivent donc pas masquer la réticence académique, commune à la sociologie et à la science politique, à participer à la production d'évaluations. Cette réticence ne saurait être attribuée au rapport à la quantification de ces deux disciplines ; il s'agit d'une position professionnelle de principe, fondée sur une réserve de caractère épistémologique qui les place dans une approche critique des institutions.

2.2.3. De la participation à l'évaluation à une forme d'ignorance

À la fin des années 1980, certains laboratoires de recherche se positionnèrent sur le « créneau » de l'évaluation comme prestataires potentiels auprès des services de l'État, central et local. En l'espèce, ce fut le cas du laboratoire CEOPS (Conception d'évaluation pour les organisations et les politiques publiques), créé à Lyon en 1988 par le ministère de l'Équipement, alors en pointe sur la modernisation de la gestion publique. Ce laboratoire de l'École nationale des travaux publics de l'État (ENTPE) a réuni une petite dizaine de chercheurs, principalement des politistes et des sociologues¹⁷ sous la direction d'Éric Monnier jusqu'en 1994. Le centre a notamment été à l'origine du programme Means¹⁸ qui a débouché sur un guide dont l'objectif est d'améliorer et de promouvoir les méthodes d'évaluation. Progressivement, l'équipe de recherche se mua en une équipe d'experts qui contribuaient à forger les méthodologies de la Commission européenne en matière de grands programmes territorialisés notamment. La bascule vers le conseil intervint définitivement en 1994.

D'autres centres de recherche spécialisés¹⁹ sur des domaines de politiques publiques (principalement en politiques urbaines) ou plus généralement dans le champ émergent de l'analyse des politiques publiques [notamment le Cerat (Centre de recherche sur le politique, l'administration et le territoire) à Grenoble] s'intéresseront également à l'évaluation, en réalisant quelques missions, pour ensuite s'en détourner. Plusieurs raisons peuvent expliquer ce phénomène. La relative résistance du monde universitaire à l'égard des activités évaluatives trouve son origine dans la volonté des sciences sociales de garder une certaine forme d'indépendance vis-à-vis du pouvoir

¹⁷ On note parmi eux V. Spenehauer et P. Warin, qui menaient à cette époque des travaux doctoraux sur l'évaluation.

¹⁸ « Méthodes d'évaluation des actions de nature structurelle » suscitées par la Direction générale (politique régionale et cohésion) de la Commission européenne.

¹⁹ Le répertoire des évaluateurs des politiques publiques de janvier 1990 élaboré par le Plan urbain recense plus d'une quarantaine de centres de recherche susceptibles de mener des évaluations et/ou ayant une première expérience sur le sujet.

politique. La position de conseiller du prince et/ou de prestataire de services est récusée au profit d'une forme de « neutralité axiologique » car, à l'inverse de l'évaluateur, le chercheur a l'initiative de ses questions, établit son calendrier, et se garde de porter des jugements de valeur. L'activité « à l'intérieur » de l'EPP apparaît alors d'autant plus difficilement compatible avec une position au sein de l'Université ou d'un laboratoire de recherche qu'elle n'est pas reconnue dans l'évaluation de la recherche par les pairs. L'absence de passerelles entre les activités de recherche et celles d'expertise, voire de consultance, est très bien illustrée par un politiste, directeur de recherche au CNRS qui se demande si l'on n'est pas condamné à « friser le hors-jeu de l'activité scientifique » quand on participe à des opérations d'évaluation [Warin, 2008]. Pour sa part, P. Duran invoque « la culture de désintéressement » des chercheurs et le « scepticisme exagéré des politiques à l'égard de leur production » comme autant de freins à d'éventuelles collaborations [Duran, 1996]. La querelle est d'autant plus importante que les activités d'évaluation, et plus généralement de consultance auprès d'organisations publiques ou privées, sont réputées, à tort ou à raison, financièrement rentables ; ce qui explique également pourquoi le chercheur qui s'y investit est réputé « perdre son âme ». Par ailleurs, le faible investissement des centres de recherche dans l'évaluation peut être expliqué par la difficulté pour le monde universitaire de rentrer dans les « canons » des commandes des collectivités territoriales ou autres organisations publiques²⁰. La capacité de réaction dans un temps court (alors que la recherche privilégie le travail long) ainsi que le caractère opérationnel des commandes qui exigent la production de recommandations utiles constituent des freins souvent évoqués comme des obstacles au positionnement du milieu universitaire, en dehors des économistes dont on a vu à l'instant la compatibilité de leur position académique avec la position de conseiller. Malgré ce constat, certains politistes ont insisté sur l'intérêt d'une implication académique croissante dans l'évaluation, soit parce que dictée par une nécessité empirique [Faure et Smith, 1994] soit parce que fondée sur une dimension citoyenne [Gaxie et Laborier, 2003].

3. Conclusion

S'il convient de souligner l'apport important des sciences sociales dans la construction progressive d'une « évaluation à française », celui-ci est finalement teinté d'une grande ambiguïté. Des attitudes de méfiance, voire d'ignorance, sont évidemment présentes. On peut penser que certaines sont stratégiques, en raison de concurrences dans le champ de l'évaluation. Mais ce positionnement des acteurs contribue, avec les facteurs institutionnels évoqués en première partie, à expliquer la difficile constitution d'un milieu professionnel spécifique de l'EPP en France, avec ses filières de formation, ses références éthiques et déontologiques, ou encore ses normes de qualité.

L'économie, familière des méthodes de quantification, fournit à l'EPP de nombreux participants, mais seule une petite minorité de chercheurs s'est, jusqu'à présent, intéressée à l'existence d'un espace professionnel qui aurait ses propres interrogations de méthode et de références déontologiques. La sociologie et la science politique, de leur côté, mettent en avant, dans le cas français, des obstacles épistémologiques

²⁰ *Propos d'entretien avec un ancien commanditaire d'évaluation, novembre 2006.*

pour une participation pleine et entière. Ce sont ces réticences et ces traditions académiques qui expliquent que, contrairement au cas américain ou britannique par exemple, l'EPP reste un objet largement méconnu de nos sciences sociales. Cette méconnaissance vient renforcer l'action des mécanismes institutionnels hérités de l'histoire : structuration du système politique, idéologie héritée d'un intérêt général, rôle de certains grands corps. En comparaison internationale, cette situation colore fortement le cas français. Il y a fort à parier que, contrairement aux analyses qui postulent une efficacité souterraine du triomphe du management public international, ces particularités resteront encore longtemps structurantes. Au demeurant, le fait que sociologie et science politique, en France, aient gardé leurs distances avec l'évaluation a sans doute constitué un antidote à l'influence négative de la « consultance » qu'on a pu observer dans les sciences sociales au Royaume-Uni, par exemple [Smith, 2002]. Avec les récents tournants politiques en France : politisation renouvelée de la réforme de l'État, fragilisation des arrangements institutionnels historiques en France [Bezes, 2009, p. 474], la transformation incessante du management public ne va certes pas s'arrêter, mais rien ne laisse vraiment augurer qu'une activité d'EPP parviendrait à affirmer sa spécificité de façon stable en France. Du coup, la pratique de l'évaluation pourrait cependant se « banaliser » un peu à la façon dont elle l'est déjà dans des pays comme les États-Unis, les Pays-Bas ou le Canada, parmi d'autres techniques à « visée évaluative ». L'originalité revendiquée de la démarche, fortement thématisée avec le CSE, disparaîtrait alors dans la panoplie des techniques du management public.

Bibliographie

- BARBIER J.-C. (2004), « Devising and Using Evaluation Standards: The French Paradox », in Schwartz R., Mayne J., eds, *Quality Matters, Seeking Confidence in Evaluating, Auditing and Performance Reporting*, Transaction Publishers, Rutgers, NJ, p. 21-41.
- BARBIER J.-C. (2007), « Evaluation and Political Communication in the French Public Employment Service », in R. Boyle, J. Breul and P. Dahler-Larsen, eds, *Open to the Public: Evaluation in the Public Arena*, Transaction Publishers, Rutgers, NJ, p. 21-41.
- BARBIER J.-C. (2008), « La transformation des forums de politique sociale en France depuis les années 1990 », in O. Giraud et P. Warin, dir., *Politiques publiques et démocratie*, Paris, La Découverte, p. 217-239.
- BASLE M. et MALGRANGE P. (2006), Présentation générale. Aide à la décision publique : développements récents, *Économie et Prévision*, n° 175, p. 1-6.
- BEHAGHEL L., CRÉPON B., GUITARD J. et GURGAND M. (2008), *Évaluation d'impact de l'accompagnement des demandeurs d'emploi par les opérateurs privés de placement et le programme Cap vers l'entreprise*, Rapport intermédiaire, Crest et École d'économie de Paris.
- BEZES P. (2009), *Réinventer l'État. Les réformes de l'administration française (1962-2008)*, Puf, Paris.
- BOITEUX M. (1994), *Transports : pour un meilleur choix des investissements*, Rapport du groupe Commissariat général du Plan, La Documentation française, Paris.
- BOURDIN J. (2001), *Rapport d'information fait au nom de la délégation du Sénat pour la planification sur l'information économique aux États-Unis*, Sénat, session 2000-2001, n°326, Paris, 382 p.

- CSE (Conseil scientifique de l'évaluation) (1992), *L'évaluation, de l'expertise à la responsabilité*, La Documentation française, Paris.
- DELEAU M., NIOCHE J.-P., PENZ P., POINSARD R. (1986), *Évaluer les politiques publiques : méthodes, déontologie, organisation*, La Documentation française, Paris.
- DURAN P. (1996), « Les non-dits de l'évaluation » dans G. Timsit (dir.), *Les administrations qui changent : innovations techniques ou nouvelles logiques ?*, Puf, Paris, p. 161-186.
- DURAN P., MONNIER E. (1992), « Le développement de l'évaluation en France, nécessités techniques et exigences politiques », *Revue française de science politique*, vol. 42, n° 2, p. 235-262.
- DURAN P., MONNIER E., SMITH A. (1995), « Evaluation à la française: Towards a New Relationship between Social Science and Public Action », *Evaluation*, vol. 1, n° 1, p. 45-63.
- FAURE A., SMITH A. (1994), « L'évaluation objet de savoir et enjeu de pouvoir », *Sciences de la société*, n°32, p. 149-152.
- FONTAINE J., WARIN P. (2000), « Retour d'évaluations : la régionalisation des politiques publiques entre affichage et politiques publiques », *Pôle Sud*, vol. 12, n° 1, p. 95-112.
- FURUBO J.-E., RIST R.C., STANDAHL R. (2002), *International Atlas of Evaluation*, Rutgers, NJ, Transaction Publishers.
- GAXIE D. ET LABORIER P. (2003), « Des obstacles à l'évaluation des politiques publiques et de quelques pistes pour tenter de les surmonter », in Favre G., Hayward J., Schmeil Y. (dir.), *Être gouverné, mélanges en l'honneur de Jean Leca*, Presses de la FNSP, Paris, p. 201-224.
- GREFFE X. (1997), *L'évaluation des projets publics*, Paris, Economica.
- HASSEUTEUFEL P., SMITH A. (2002), « Essoufflement ou second souffle ? L'analyse des politiques publiques « à la française » », *Revue française de science politique*, vol. 52, n° 1, p. 53-73.
- IHL O., KALUSZYNSKI M., POLLET G., (dir.) (2003), *Les sciences de gouvernement*, Economica, Paris.
- JACOB S. (2005), « La volonté des acteurs et le poids des structures dans l'institutionnalisation de l'évaluation des politiques publiques (France, Belgique, Suisse et Pays-Bas) », *Revue française de science politique*, vol. 55, nos 5-6, p. 835-864.
- JOBERT B. (1992), « Représentations sociales, controverses et débats dans la conduite des politiques publiques », *Revue française de science politique*, vol. 42, n° 2, p. 219-234.
- JOBERT B. ed, (1994), *Le tournant néo-libéral en Europe*, L'Harmattan, Paris.
- LASCOUMES P., SETBON M. (1996), « L'évaluation pluraliste des politiques publiques : enjeux, pratiques, produits », Note pour le Commissariat Général du Plan, miméo, Paris, 17 p.
- LASCOUMES P., LE GALÈS P. (2004), *Gouverner par les instruments*, Presses de la Fondation nationale des sciences politiques, Paris.
- LECA J. (1993), « Sur le rôle de la connaissance dans la modernisation de l'État et le statut de l'évaluation », *Revue française d'administration publique*, n° 66, p. 185-196.
- LECA J. (1997), « L'évaluation comme intervention, sciences sociales et pratiques administratives et politiques », in *L'évaluation des politiques publiques*, Actes du colloque GAPP-CNRS-ENS Cachan des 5 et 6 février, miméo, 19 p.

- LECA J., THOENIG J.-C. (1997), « Évaluer pour gouverner autrement », *Le Monde* du 9 décembre.
- MATYJASIK N., MÉASSON L. (2007), « L'évaluation entre opportunité et effet » in Faure A., Négrier E., dir, *Les politiques publiques à l'épreuve de l'action locale : critiques de la territorialisation*, L'Harmattan, Paris, coll. « Questions contemporaines », p. 170-176.
- MÉNY Y., THOENIG J.-C. (1989), *Politiques publiques*, Puf, Paris.
- MONNIER E., (1992), *Évaluations de l'action des pouvoirs publics*, 2^e éd., Economica, Paris [1^{re} éd. 1987].
- MULLER P. (2000), « L'analyse cognitive des politiques publiques : vers une sociologie politique de l'action publique », *Revue française de science politique*, vol. 50, n° 2, p. 189-207.
- NIOCHE J.-P. (1982), « De l'évaluation à l'analyse des politiques publiques », *Revue française de science politique*, vol 32, n° 1, p. 32-61.
- PERRET B. (2001), *L'évaluation des politiques publiques*, La Découverte, Paris, coll. « Repères ».
- PRESSMAN G., WILDAVSKY A. (1973), *Implementation*, University of California Press, Berkeley.
- Sabatier P.A. (1998), « The Political Context of Evaluation Research. An Advocacy Coalition Perspective », in Kessler M.C., Lascoumes P., Setbon M., Thoenig J.C. (dir.), *Évaluation des politiques publiques*, L'Harmattan, Paris, p. 129-146.
- SPENLEHAUER V. (1998), *L'évaluation des politiques publiques, avatar de la planification*, Thèse de doctorat de science politique, Grenoble, 586 p.
- SPENLEHAUER V. (2006), « Revue critique de Steve Jacob, *Institutionnaliser l'évaluation des politiques publiques. Études comparées des dispositifs en Belgique, en France, en Suisse et aux Pays-Bas* », *Politiques et Management Public*, n° 2, p. 162-165.
- SPENLEHAUER V., WARIN P., (2000), « L'évaluation au service des conseils régionaux », *Sociologie du travail*, vol. 42, n° 2, p. 245-262.
- SMITH A. (2002), « Grandeur et décadence de l'analyse britannique des politiques publiques » *Revue française de science politique*, vol. 52, n° 1, p. 23-35.
- THOENIG J.-C. (1976), *Enseigner le management public. Le courant actuel aux États-Unis*, Enseignement et Gestion, n°15, FNEGE, Paris.
- THOENIG J.-C. (2005), « Pour une épistémologie des recherches sur l'action publique », in D. Filâtre et G. de Terssac, (dir.), *Les dynamiques intermédiaires au cœur de l'action publique*, Toulouse, Octarès, p. 285-306.
- TOULEMONDE J. (2000), « Evaluation culture(s) in Europe: Differences and Convergence between National Policies », *Vierteljahrshefte zur Wissenschaftsforschung*, 3, p. 350-357.
- VATIN F. (2008), « L'esprit d'ingénieur : pensée calculatoire et éthique économique », *Revue Française de Socio-Économie*, n°1, p. 131-152.
- VIVERET P. (1989), *L'évaluation des politiques et des actions publiques*, Rapport au Premier ministre, Paris, La Documentation française.
- WARIN P. (2008), « Une recherche publique dans le *policy process* : tactique et raison scientifique », *Les nouveaux acteurs politiques*, communication au Colloque de la Société Québécoise de Science Politique, Université de Montréal.
- WEISS C.H., MURPHY-GRAHAM E., PETROSINO A. AND GANDHI A.G. (2008), « The Fairy Godmother and Her Warts: Making the Dream of Evidence-Based Policy Come True », *American Journal of Evaluation*, vol. 29, n° 1, p. 29-47.