



HAL
open science

Les scènes multiples de l'évaluation : Les problèmes récurrents de son institutionnalisation

Clément Lacouette-Fougère, Pierre Lascoumes

► To cite this version:

Clément Lacouette-Fougère, Pierre Lascoumes. Les scènes multiples de l'évaluation : Les problèmes récurrents de son institutionnalisation. [Rapport de recherche] Laboratory for Interdisciplinary Evaluation of Public Policies. 2013, pp.120. hal-02408128

HAL Id: hal-02408128

<https://sciencespo.hal.science/hal-02408128v1>

Submitted on 12 Dec 2019

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.



SciencesPo.

LIEPP

Laboratoire interdisciplinaire d'évaluation des politiques publiques

A photograph of the interior of a grand, circular building, likely the SciencesPo courtyard. The building has multiple levels of windows and a series of columns at the bottom, each with a red curtain. A person is visible in the distance, standing near one of the columns. The lighting is warm, highlighting the architectural details.

« Les scènes multiples de l'évaluation »
Les problèmes récurrents de son institutionnalisation

Clément Lacouette-Fougère
Pierre Lascoumes

Policy Paper n°1 - mai 2013

www.sciencespo.fr/liepp

Les scènes multiples de l'évaluation
Les problèmes récurrents de son
institutionnalisation

© Sciences Po / LIEPP 2013

Directeurs de publication: Etienne Wasmer & Cornelia Woll

Maquette: Alexandre Biotteau & Chloé Touzet

Photo couverture: CC-BY-SA Malouette

Sciences Po – LIEPP, 27 rue Saint Guillaume, 75007 Paris - France

Tél : +33(0)1.45.49.83.61

www.sciencespo.fr/liepp

Les scènes multiples de l'évaluation
Les problèmes récurrents de son institutionnalisation

Clément Lacouette-Fougère¹
Pierre Lascoumes²

mai 2013

¹ (Sciences Po, Planète Publique), clement.lacouette@sciences-po.org

² (Sciences Po, CEE/LIEPP), pierre.lascoumes@sciences-po.fr

L'objectif du LIEPP :
Étudier et évaluer les politiques publiques dans
une perspective interdisciplinaire

Un laboratoire d'excellence (labex)

Le Laboratoire interdisciplinaire d'évaluation des politiques publiques (LIEPP) est un laboratoire d'excellence (Labex), distingué par le jury scientifique international désigné par l'Agence nationale de la recherche (ANR) et financé pour 10 ans dans le cadre des investissements d'avenir.

Une équipe pluridisciplinaire

Ce projet réunit plus de cinquante chercheurs et des professeurs de Sciences Po venant du Département d'économie, du Centre d'études européennes (CEE), du Centre de sociologie des organisations (CSO) et de l'Observatoire sociologique du changement (OSC) ainsi que des juristes et des historiens. Porté par Sciences Po, le LIEPP associe également des chercheurs d'autres universités sur des projets d'évaluation et s'appuie sur les coopérations scientifiques à l'international des centres de recherche fondateurs.

Des thématiques prioritaires

Le LIEPP repose sur une méthode innovante, fondée sur la pluridisciplinarité et la confrontation d'évaluations qualitatives, comparatives et quantitatives d'une même politique. L'activité actuelle porte sur plusieurs thèmes, déjà présents à Sciences Po en matière d'évaluation des politiques publiques : méthodes d'évaluation, éducation, enseignement supérieur, politiques d'emploi, politiques sociales et fiscales, politiques de santé.

Sciences Po / LIEPP - 27, rue Saint-Guillaume - 75007 Paris France T/ +33 (0)1 45 49 83 61

www.sciencespo.fr/liepp

« Toutes ces considérations s'appuient sur une histoire assez courte, et ne valent pas évaluation ! Elles le deviendront pourtant si l'évaluation connaît le sort de nombre de réformes administratives, c'est à dire d'être mises en sommeil avant que d'avoir eu le temps de se développer ».

Sylvie Trosa, Conseil scientifique de l'évaluation, Cerisy, juin 1992.

Table des matières

Introduction	9
PARTIE 1 : Les enjeux de l'évaluation de l'action publique à travers les débats initiaux	19
1.1. <i>L'évaluation comme légitimation et comme contre pouvoir</i>	19
1.2. <i>Les questions de modernisation de l'action publique et de compréhension des administrations</i>	28
1.3. <i>Le renforcement des secteurs</i>	35
1.4. <i>Les approches démocratiques : de l'évaluation pluraliste à l'évaluation émancipatrice</i>	50
1.5. <i>Sortie des controverses et évaluation technologique et scientifique</i>	56
1.6. <i>Le retour de la performance publique</i>	62
PARTIE 2 : L'organisation, les pouvoirs, les finalités	71
2.1. <i>L'évolution de toutes les inspections et contrôle des comptes vers l'évaluation</i>	71
2.2. <i>Les Parlements</i>	80
2.3. <i>Les activités d'évaluation au niveau territorial</i>	92
2.4. <i>Autre acteurs-autres rattachements</i>	106
2.5. <i>L'évaluation comme expertise : la place du secteur privé</i>	111
CONCLUSION : bilan de l'institutionnalisation	127
<i>Des scènes « ouvertes »</i>	127
<i>Six lignes de débat</i>	130
<i>« Tous compétents ! » : L'indéfinissable approche à la Française</i>	133
<i>Controverses et résistances autour du mythe de l'évaluation</i>	138
Bibliographie	147

Table des encarts

Encart n° 1 : Résumé de la thèse de Vincent Spenlehauer	24
Encart n° 2 : Le Comité national d'évaluation (CNE).....	41
Encart n° 3 : Le « Policy Planning Staff » aux Etats-Unis	47
Encart n° 4 : Les quatre générations d'évaluation selon Guba et Lincoln (1989) :.....	51
Encart n° 5 : Les pratiques d'évaluation du canton de Genève.....	52
Encart n° 6 : Les évaluations émancipatrices, une 5ème génération d'évaluations.....	56
Encart n° 7 : L'institutionnalisation du <i>Technological Assessment</i> en Europe.....	59
Encart n° 8 : What is a value for money study ?.....	73
Encart n° 9 : Le Club Cambon pour une Cour des Comptes évaluative	75
Encart n° 10 : Une nouvelle mission de conseil.....	78
Encart n° 11 : Rapport sur la politique vaccinale de la France (février 2013).....	80
Encart n° 12 : Loi de 1991 sur les déchets nucléaires.....	84
Encart n° 13 : L'évaluation des politiques publiques à l'Assemblée nationale.....	88
Encart n° 14 : Enquête de 2002 sur les démarches d'évaluation dans les collectivités territoriales.....	103
Encart n° 15 : L'évaluation de la politique de la ville en France.....	106
Encart n° 16 : Les cabinets de conseil en évaluation en France.....	116
Encart n° 17 : Le titre d'évaluateur accrédité (ÉA) au Canada	117
Encart n° 18 : Les neuf compétences génériques des évaluateurs	118
Encart n° 19 : Le bilan des prestations confiées aux cabinets privés dans la RGPP.....	125
Encart n° 20 : Tentative (manquée) de putsch parlementaire contre la rue Cambon (juillet 2008).....	136
Encart n° 21 : L'idéologie et la tyrannie de l'évaluation.....	139
Encart n° 22 : Psychothérapies : le rapport d'évaluation censuré	144

Introduction

Durant la période au cours de laquelle a été préparé ce « *position paper* »³, l'évaluation des politiques publiques a été, une fois encore, à l'actualité. La nomination en septembre 2012 de Madame Yannick Moreau⁴ à la tête d'une « instance de réflexion prospective et d'expertise sur les politiques publiques » montre l'intérêt du gouvernement actuel pour les questions d'évaluation. Plus exactement, cette décision atteste du caractère récurrent d'un besoin d'analyse de la rationalité de l'action publique et de ses effets, et cela depuis, au moins ... une soixantaine d'années.

L'objet de cette note est de rendre compte des débats et des multiples propositions faites en France, mais aussi dans les principales démocraties sur l'institutionnalisation des activités d'évaluation. Il s'agit de dégager les principes et les justifications qui ont accompagné le développement de l'évaluation et de caractériser les choix organisationnels effectués en vue de son opérationnalisation.

L'évaluation, un « marronnier » de l'action gouvernementale

On peut avancer que l'évaluation est un « marronnier » de l'action gouvernementale, un passage obligé sous-tendu par deux dynamiques, plus ou moins bien articulées, et très variables dans le temps : d'un côté, la modernisation de l'Etat, de ses structures et de ses moyens ; d'un autre côté, la recherche d'une plus grande efficacité de ses actions (en fait, le plus souvent une réduction de ses coûts). En France, l'alternance politique du printemps 2012 a conduit à écarter une « Revue générale des politiques publiques » (RGPP) jugée réductrice, pour la remplacer par une « Modernisation de l'action publique » (MAP) dont le contenu précis reste à définir.

³ Ce travail a été réalisé de mai 2012 à mars 2013. Il s'appuie sur l'exploitation de la littérature produite depuis les premières années de mise en œuvre de l'évaluation à l'étranger (années 50-60) et en France (années 70-80) jusqu'aux développements les plus récents. La liste complète des références bibliographiques mobilisées est présentée à la fin de la note.

⁴ Conseillère d'Etat, en charge du Haut conseil du dialogue social.

Amorcée dans les années 1960-1970, l'évaluation des politiques publiques a pris de l'importance en voyant son assise institutionnelle renforcée à partir des années 90. Mais le parcours de cette activité combinant la production de connaissance et les choix politiques reste aléatoire. Il est aussi peu cumulatif dans son ensemble. Cinq phénomènes observés durant ces huit derniers mois attestent des forces en présence et des jeux de rôles attendus en ce domaine.

1 - Madame Moreau chargée par le premier ministre d'imaginer « un lieu de dialogue, de réflexion prospective et d'expertise sur les politiques publiques ouvert à l'ensemble des acteurs sociaux » a rendu un rapport le 5 décembre 2012⁵. Elle propose la création d'un Commissariat général à la stratégie et à la prospective (CGSP) qui serait « un lieu transversal de concertation et de réflexion 'à froid', c'est-à-dire un lieu de synthèse et de partage, au service de la construction d'une vision de l'avenir ». Beaucoup d'observateurs ont vu dans cette proposition une réactivation du Commissariat général du Plan, remplacé en 2005 par le Centre d'analyse stratégique » (CAS)⁶. A travers cette initiative, se lit la première constante française : la relation historique qui unit évaluation et prospective. Cette superposition d'objectifs divergents a pour conséquence, selon plusieurs analystes (Spencehauer, 1998), d'aliéner la première à la seconde comme nous le rappellerons ultérieurement. L'articulation entre évaluation et prospective est une fausse évidence. Il s'agit, fondamentalement, d'exercices relevant de logiques différentes.

Autre nouveauté interministérielle, celle introduite en avril 2013 par le Secrétariat général pour la modernisation de l'action publique (SGMAP). Il a créé un « Laboratoire de l'innovation publique » afin d' « esquisser les politiques publiques de demain au plus près des attentes des usagers ». L'objectif est de « faire remonter les expériences innovantes des collectivités locales et des services déconcentrés, de pratiquer des expérimentations à petite échelle et de diffuser des modèles. Ce projet reprend une pratique danoise le « MindLab » qui mène une réflexion interministérielle sur l'innovation publique depuis une dizaine d'années.

⁵ Elle avait constitué un groupe de travail réunissant Ph. Aghion, M. Guillou, P. Rosanvallon, H. Rouilleaut et L. Schweizer.

⁶ M. Noblecourt, « Le fantôme du Plan est de retour », *Le Monde*, 04/02/2013.

2 - Lors de l'audience solennelle de rentrée de la Cour des comptes, le 10 janvier 2013, son président Didier Migaud a déclaré qu'une évaluation plus régulière des politiques publiques était une nécessité. « Il s'agit de soumettre les politiques publiques à des évaluations plus systématiques et plus régulières. Cet exercice « qu'on le nomme RGPP ou MAP est essentiel pour le bon fonctionnement de l'Etat et d'une démocratie ». A cette occasion, il a repris un diagnostic régulièrement formulé depuis une quarantaine d'années concernant la faiblesse de la culture de l'évaluation dans les institutions publiques françaises et les obstacles bien connus dus à l'éparpillement des opérateurs : « Il en est résulté des sédimentations complexes de dispositifs insuffisamment évalués ». Le contexte de volonté de réduction des déficits publics incite au renforcement de ces activités. Il s'agit là de la deuxième constante française : le rôle leader de la Cour des comptes en matière d'évaluation. Son évolution du contrôle de la régularité à l'analyse de l'action publique a été amorcée dans les années 80 et fortement légitimée par la réforme constitutionnelle de 2008. La Cour des comptes, comme dans d'autres pays, mais avec de plus faibles moyens, reste l'acteur clef de l'évaluation des politiques publiques. Cela sera démontré dans la deuxième partie de cette note.

3 - En décembre 2012, trois inspections générales (de l'Administration, des Finances et des Affaires sociales) ont publié un « Cadrage méthodologique de l'évaluation des politiques partenariales » qui fait suite au séminaire gouvernemental du 1^{er} octobre 2012 relatif à la modernisation de l'action publique durant lequel il a été décidé d'engager « un travail d'évaluation des politiques publiques partenariales avec l'ensemble des acteurs concernés (...) pour construire une vision collective des enjeux, des finalités et des modalités de mise en œuvre de chaque politique publique ».

Le guide s'appuie sur les expériences pratiques des diverses inspections, ainsi que sur les démarches développées par une série d'organismes français (le Conseil d'orientation des retraites, la préparation du Livre blanc de la défense, les travaux préparatoires à la Conférence nationale de lutte contre la pauvreté, etc.) et étrangers (« *Programm reviews* » canadiens). Le guide pose deux grandes séries de recommandations. La première porte sur l'établissement de diagnostics ; la seconde concerne l'élaboration de scénarios de réforme. La troisième constante française repose sur le fait que

l'essentiel des activités d'évaluation est mené au sein des secteurs administratifs. Les Inspections y tiennent un rôle important mais non décisif. L'évaluation de l'action publique a été développée de façon très sectorisée, dans la plus grande hétérogénéité. Les échanges de méthodes et l'apprentissage de bonnes pratiques (telles le guide de 2013) sont l'exception, comme nous le montrerons dans la première partie. Chaque secteur administratif a développé de façon cloisonnée son approche de l'évaluation et ils la pratiquent davantage selon ses logiques internes et ses routines qu'en fonction de modèles de référence partagés.

4 - Pour compléter le tableau, regardons ce qu'il en est des activités du Parlement en ce domaine. Une réforme de 2009 a spécifié cette forme de régulation politique de l'action publique. Une semaine par mois, dite de « contrôle », les députés ont l'opportunité de débattre de sujets qui ont, en théorie, fait l'objet d'une évaluation antérieure. Chaque groupe peut ainsi présenter son analyse politique d'un domaine et la délibération collective est censée fournir des éléments de réflexion en vue de décisions ultérieures. Si démocratiquement l'initiative est intéressante, elle ne suscite qu'un très faible intérêt chez les élus, faute de vote à accomplir et faute de possibilité d'une valorisation individuelle (à l'inverse des questions au gouvernement). Le débat qui a eu lieu le 23 février 2013 sur le budget de l'Union européenne – sujet que l'on pourrait croire sensible – n'a fait que confirmer la désaffection des élus pour ce type d'exercice collectif. Au mieux une vingtaine de députés étaient en séance, chacun lisant à tour de rôle un discours pré-rédigé et aucun débat n'a été engagé. Comme l'indiquent des assistants parlementaires sur leur blog : « *Ces débats sont ennuyeux, c'est du bla-bla dont on pourrait faire l'économie ...Le cœur du problème est là : l'évaluation et le contrôle ne 'payent' pas politiquement, car cela demande beaucoup de travail pour, souvent, une très faible visibilité médiatique* »⁷. C'est la quatrième constante française : contrairement à d'autres pays, le Parlement, malgré diverses tentatives n'est jamais parvenu à s'imposer ni comme un acteur clef de l'évaluation, ni même comme un acteur réellement mobilisé sur le sujet.

⁷ S. Le Goff, « Les missions de contrôle et d'évaluation », *Les cuisines de l'Assemblée*, 25 février 2015. <http://blogs.lexpress.fr/cuisines-assemblee/>

5 – Enfin, qu'en est-il du côté des activités d'évaluation menées par le secteur privé des consultants ? Ce cinquième élément permet de boucler en quelques lignes un panorama des principaux acteurs publics de l'évaluation. En juillet 2012, deux étudiants publient une tribune dans *Le Monde* intitulée « Politique publique, pour une dépense efficace »⁸ qui a déclenché d'intéressantes réactions dans le petit monde des praticiens de l'évaluation. En effet, avec une innocence rafraîchissante ces deux jeunes chercheurs se font les avocats du renforcement des pratiques d'évaluation, en particulier sous l'angle des démarches et des méthodes. Prenant appui sur la théorie économique, ils développent l'exemple de la politique d'aide au logement pour montrer la « complexité » des choix d'action publique et l'ambiguïté, voire l'ambivalence, de bien des programmes. Sans le savoir vraisemblablement, ils reprennent les antennes les plus rabâchées du domaine : « Sans être une solution miracle, l'évaluation rigoureuse des politiques publiques permet de renforcer l'efficacité de l'Etat-providence. C'est en justifiant au contribuable l'allocation des ressources publiques et les résultats obtenus, que le lien entre citoyen et Etat sera préservé ».

Ils soulignent les insuffisances de l'institutionnalisation de l'évaluation en France et son manque d'autonomie (Cour des comptes, inspections générales, Assemblée nationale) qui « ne sont pas à même de réaliser ce type d'étude. L'administration peut auditer une politique publique, elle peut piloter sa mise en œuvre, mais l'évaluation est une tâche radicalement différente, qui demande des outils techniques particuliers ». Ils prônent une approche pluri-disciplinaire en valorisant l'évaluation expérimentale du RSA qu'ils présentent comme une nouveauté, en toute méconnaissance des travaux multiples ayant accompagné la mise en œuvre du RMI. Ils soulignent l'intérêt des engagements universitaires dans ce type d'activités (LIEPP, IDEP, IPP)⁹ et appellent à la réalisation d'une charte commune « afin d'assurer la cohérence et la pertinence des évaluations produites : une charte commune ». Ces jeunes Messieurs Jourdain formulent aussi une proposition radicale : « il serait intéressant d'associer administrations publiques, régulateurs et monde académique dans l'évaluation des grandes réformes économiques et sociales » ... ! Ils concluent en rappelant les

⁸ V. Hirsch (Sciences-po et ENSAE) et Q. Lafay (ENS), *Le Monde*, 9/07/2012

⁹ LIEPP : Laboratoire interdisciplinaire d'évaluation des politiques publique, Sciences-Po Paris, IDEP : Institut d'économie publique, Aix-Marseille, IPP : Institut de politiques publiques, Ecole d'économie de Paris.

réticences (bien connues) des acteurs politiques à l'évaluation et la nécessité de faire émerger « une réelle culture de l'évaluation », tant par « un changement de mentalités que l'adaptation des dispositifs légaux actuels ».

Tous ceux qui depuis deux ou trois décennies se sont évertués à amorcer le programme ci-dessus (présenté avec autant de conviction que d'idéalisme) se sont un peu étranglés. Le premier texte a reçu quelques jours plus tard une réplique polie mais ferme de neuf responsables historiques et actuels de la Société française de l'évaluation (SFE) qui regroupe à peu près six cents praticiens¹⁰. Le texte est très posé et ne reflète pas les propos ironiques que la prise de position des jeunes profanes avait immédiatement suscitée dans le milieu. Il rappelle l'antériorité des pratiques, en particulier à partir des années 90. Il présente les termes de la Charte publiée presque dix ans plus tôt en 2003 : « L'appel des deux jeunes étudiants (...) est une invitation bienvenue à en développer l'audience et en enrichir le contenu ». La réponse de la SFE rappelle les principes méthodologiques qu'elle promeut et l'importance des activités de sensibilisation et de formation qu'elle mène depuis plus de dix ans. Enfin, très consensuellement, le texte conclut en appelant à une action collective : « afin de faire reconnaître l'apport de l'évaluation des politiques publiques (...) la SFE appuie l'idée d'un *Grenelle de l'évaluation* réunissant tous les acteurs concernés ».

Ces échanges sont significatifs d'une cinquième caractéristique de la situation française, celle de la méconnaissance du domaine et des pratiques de l'évaluation. Faute d'institutionnalisation forte, il y a toujours de la place pour les Messieurs Jourdain découvrant avec aplomb ce qui se dit et s'écrit depuis des décennies. Et cela aussi bien dans les secteurs politiques, administratifs qu'universitaires. En France, il est aisé de s'auto-proclamer spécialiste de l'évaluation en toute méconnaissance de l'analyse de l'action publique (sous secteur de la science politique qui a cinquante ans de travaux à son actif), et en toute ignorance des multiples démarches méthodologiques expérimentées sur des terrains très diversifiés depuis des décennies. L'humble réplique de la société professionnelle (SFE) qui est la seule organisation à s'être préoccupée d'enrichir ses méthodes par des échanges d'expériences internationaux et par une réflexion collective régulière est

¹⁰ Barbier J.C., Baslé M., Delarue J., Matyjasik N., Mouterde F., Outin J.L., Perret B., Toulemonde J., « Développer et enrichir l'évaluation des politiques et des programmes publics », *Le Monde*, 31/07/2012.

révélatrice de la faible légitimité de cette organisation face à l'arrogance des grands corps et des grandes institutions. Il est vrai que la SFE rassemble beaucoup de chercheurs et de consultants travaillant dans des bureaux d'étude privés. Ce sont eux qui composent l'essentiel des troupes de « chargés de d'étude » auxquels les diverses instances publiques ont recours pour réaliser les démarches concrètes d'évaluation, sur la base d'appel d'offres ou de gré à gré.

Etudier l'institutionnalisation

Cette étude a pour but d'opérer un retour analytique sur les principales étapes et choix théoriques qui ont accompagné la conceptualisation et l'institutionnalisation des actions d'évaluation des politiques et programmes publics afin de mettre en perspective les débats actuels sur la place de cette activité en France. Peu de travaux antérieurs fournissent des points de repère (Furubo, Rist, Sandahl, 2002). Le principal est celui de S. Jacob et F. Varone (2004) qui distingue les niveaux d'institutionnalisation à partir de quatre grandes variables : la pratique effective (fréquence des appels politiques, part du budget consacrée à ce type d'activité...); la présence ou non d'« incitants institutionnels » (par exemple des clauses de lois rendant obligatoire l'évaluation d'un programme d'action); les caractéristiques des structures organisationnelles (instances isolées ou dispositif inter-organisationnel); enfin, le degré de formalisation de « communautés épistémiques » (réseaux professionnels, sociétés nationales, standards de qualité, moyens d'information spécialisés).

Etudier l'institutionnalisation c'est analyser les relations entre deux grands types de variables :

- d'une part les idées, les représentations, les catégories cognitives à partir desquelles est pensée l'évaluation. Les conceptualisations sont hétérogènes, mais chacune s'efforce de donner des buts, une cohérence et une légitimité à cette activité. Chacune a aussi l'ambition de résoudre la tension entre une rigueur scientifique et un jugement politique. Nous verrons que la question du degré de dépendance/indépendance des structures est un enjeu majeur qui se matérialise dans le choix des structures : un rattachement à l'exécutif /Parlement ou une agence indépendante.

- d'autre part, l'institutionnalisation conduit à analyser les formes d'organisation, les pouvoirs et les compétences des structures, les moyens et les statuts des agents et leur degré de professionnalisation. Les solutions organisationnelles sont extrêmement variées, et la réalisation d'un panorama est en soi un projet très ambitieux (ce ne sera pas le nôtre).

L'institutionnalisation est à envisager comme une dynamique, elle est un résultat et pas un donné stable et cohérent (Lagroye et al. 2002) : « L'institutionnalisation résulte de processus qu'on ne saurait réduire à la codification des dispositifs et des règles, à la production de textes législatifs et de règles, à l'adoption d'un nouveau vocable... Elle officialise et donne en modèle des pratiques et des savoir-faire d'échanges entre acteurs, comme par exemple l'utilisation des mêmes catégories par les agents de l'Etat et les associations » (p.530). Non seulement les composantes de l'évaluation sont sujettes à des variations selon les conjonctures, mais les frontières et les identités de ces structures elles-mêmes connaissent des déplacements, des reformulations régulières. C'est pourquoi, il n'est jamais simple de répondre à des questions de fausse évidence comme : « Qu'est-ce que l'évaluation ? », « Quels acteurs la produisent ? », « Quelle finalité poursuivent-ils ? ». Analyser l'institutionnalisation, ses formes et ses degrés est une façon de répondre aux énigmes posées par la production d'une connaissance finalisée pour l'action. Cette approche tend à substituer aux questions précédentes celle-ci : « Qu'est-ce qui fait évaluation, pour qui, et comment ? »

Nous le verrons dans les pages suivantes, des questions incontournables se posent dès qu'il est question d'évaluation :

- quelles sont les finalités poursuivies ?
- comment concilier rigueur méthodologique et commande politique ?
- comment assurer la cohérence d'un projet évaluatif sans opérer des réductions appauvrissantes ?
- comment s'assurer de la qualité de la réception et des usages de l'évaluation sans être entravé par un mélange de positions et de responsabilités paralysant ?

L'analyse de cette activité de production de connaissance finalisée sous l'angle des institutions qui la revendiquent permet d'avancer dans cette réflexion. L'exercice pousse aussi à la modestie en écartant le mythe rassurant des solutions toutes faites. Beaucoup de formules ont été mises à l'œuvre, aucune n'a fait l'unanimité. Le type de commanditaire, de programme envisagé, l'espace territorial concerné et le but attribué à la connaissance attendue influent de façon déterminante sur le mode d'évaluation à mobiliser.

Une bonne appréhension de la diversité des modalités d'institutionnalisation permet aussi de consolider la légitimité de cette activité en fournissant des critères de choix tant pour la définition du type d'évaluation que l'on envisage, que pour fixer les modalités pratiques de sa mise en œuvre.

Cette étude procède en trois temps :

- tout d'abord, nous rappellerons les principaux enjeux de l'évaluation de l'action publique tels qu'ils ont été formulés lors des débats initiaux,
- ensuite, nous présenterons les principales formes d'organisation, de pouvoirs et de finalités donnés aux différents organismes en charge de cette activité,
- enfin, nous conclurons en présentant les grandes lignes d'évolution du domaine et les controverses en cours.

PARTIE 1 : Les enjeux de l'évaluation de l'action publique à travers les débats initiaux

« On doit donc être attentif au risque d'une approche purement « managériale » de l'évaluation : dans le secteur privé, comme dans le secteur public, l'évaluation est souvent une ressource de puissance pour les décideurs. ... L'évaluation devient ainsi un instrument de pouvoir supplémentaire au lieu d'être un moyen de partage du pouvoir et de savoir dans l'entreprise comme dans la cité » (P. Viveret, 1989, p.27)

1.1. L'évaluation comme légitimation et comme contre pouvoir

L'institutionnalisation de l'évaluation est souvent assimilée aux besoins de rationalisation de l'action publique tels qu'ils ont été formulés et mis en œuvre à partir des années 50 en particulier dans les pays anglo-saxons. Mais cette dynamique en recouvre deux autres, plus discrètes et néanmoins fondamentales : celle de l'organisation d'un contre-pouvoir face aux décisions de l'exécutif et celle d'une légitimation de l'intervention de l'Etat. « La naissance d'une rationalité technocratique qui deviendra dominante dans la haute administration des pays développés (s'est trouvée) confortée par les crises périodiques de légitimité auxquels auront à faire face les Etats. L'exemple américain est ici particulièrement éclairant puisque le développement des évaluations à caractère « scientifique » épouse les moments historiques de ces crises » (Viveret, 1989, p.40).

Etats-Unis : un contre pouvoir pour le législatif

Le contexte des suites de la crise de 1929 aux Etats-Unis et les programmes formant les deux « *New Deal* » (1933-1935, 1935-1938) du président Fr. D. Roosevelt conduisirent le gouvernement à réaliser des investissements inédits et la mise à disposition de budgets importants pour diverses agences gouvernementales. Les notions formulées par Laswell dès 1930 sous le concept de « *policy analysis* » reçurent dès lors une forte audience poli-

tique. Si cet auteur prônait une action publique efficace, et si possible efficiente, ce n'était pas pour réduire les budgets publics, mais parce qu'il considérait que c'était une condition essentielle de son acceptabilité tant politique (par le Congrès et les partis politiques) que sociale (soutien de l'opinion et des groupes intermédiaires). Cette période est marquée par un changement important dans le mode gouvernement. La légitimité reconnue aux Etats démocratiques n'est plus seulement à base légale et institutionnelle, elle comporte désormais une part croissante de démonstration de leur utilité et de leur efficacité.

Ce type de raisonnement a été accentué dans l'après-guerre avec l'effort de reconstruction. L'accroissement de la richesse nationale est devenu la rationalité dominante dans l'action gouvernementale. L'application systématique des critères de valeur économique qui avaient été généralisés en entreprise, concerne désormais l'Etat. C'est la période de développement du Management public sous forme des procédures de planification et de méthodes d'aide à la décision. Dans la plupart des démocraties industrialisées, le dirigisme désordonné qui avait prévalu jusqu'à l'avant-guerre fut révisé au profit d'un interventionnisme public renforcé. Mais les autorités politiques et administratives s'efforcent de rendre ces investissements légitimes en faisant la démonstration empirique de leur efficacité.

L'échec de la guerre de Corée en 1953, puis le lancement du premier spoutnik soviétique en 1957 vont conduire à une réflexion radicale sur l'organisation militaire et ses besoins en nouveaux armements. C'est dans la période qui va du début des années 50 à 1965, qu'aux USA sont développés, en particulier par la *Rand corporation* et en relation avec des universitaires, un ensemble de méthodes d'analyse micro-économique des choix publics regroupées sous l'intitulé: *Planning programming budgeting system* (PPBS). Leur principale application concerne les budgets militaires (Chelimsky, 1985). En utilisant une formule anachronique pour l'époque, on peut les qualifier « d'évaluation ex-ante ». Mais il s'agissait principalement d'aide à la planification par la tentative de détermination des coûts de recherche et de production de nouveaux armements et l'anticipation des effets potentiels de ces programmes. Le Congrès accorda une grande attention à ces travaux en vue d'exercer un contrôle sur les demandes de l'exécutif dans une période où les rapports Est-Ouest incitaient à des investissements massifs. Ainsi les pro-

jets de satellites d'observation militaire (et à terme de mise en orbite d'armements) reposaient sur des budgets colossaux mais qui s'avérèrent très flous à l'analyse. Le recours au PPBS permettra ainsi au Congrès d'en limiter de façon significative les ambitions (Hitch, Mc Kean, 1960).

Au début des années 60, c'est la « crise sociale » qui va générer indirectement une nouvelle étape dans les activités d'évaluation. Les questions raciales et les problèmes multi-dimensionnels liés à la déstructuration des centres ville et la formation des ghettos urbains acquièrent une forte visibilité aux Etats-Unis à l'occasion d'une série d'émeutes inédites. En réponse, le président L. Johnson élabore à partir de 1964 un vaste ensemble de programmes connus sous les noms de « *Great society* » et de la « *War on poverty* ». Mais l'importance du coût des programmes sociaux engagés suscita beaucoup de critiques qui incitèrent à la mise en place d'un accompagnement nettement évaluatif des décisions gouvernementales. Deux ans plus tard, les démocrates ayant perdu les élections de mi-mandat, le Congrès se mobilisa très fortement sur le suivi de ces programmes, non seulement en termes de coûts, mais aussi d'effets sociaux. Des sommes considérables ont été allouées à différents grands programmes : lutte contre les ghettos et réhabilitation des centres urbains, alphabétisation et accès au système scolaire, formation professionnelle et accès à l'emploi, lutte contre l'usage de l'alcool et des drogues et contre la mortalité infantile et les maladies vénériennes, enfin, sécurité publique et lutte contre la délinquance, en particulier celle des jeunes. Tous ces programmes d'aide furent l'objet de nombreuses investigations et cela jusqu'en 1974. Durant cette période, il y eut aux USA une généralisation des démarches d'évaluation de la performance publique entendue sous un angle à la fois budgétaire, mais aussi de mesure de l'efficacité. L'enjeu principal était de pouvoir justifier aussi bien économiquement que politiquement l'usage des fonds fédéraux.

Ces évaluations reposaient sur des « méthodologies mixtes » combinant des techniques quantitatives et qualitatives. Menées à grande échelle (et souvent comparative entre territoires et entre programmes) ces investigations mobilisèrent les universités et les bureaux d'étude privés (cf. 2.5). Aux côtés des juristes et économistes, les sociologues et anthropologues prirent une place importante à partir de 1970. Ils développèrent des approches dites

«compréhensives» enrichissant les approches économiques qui prévalaient jusqu'alors. Ils effectuèrent des analyses détaillées des systèmes d'acteurs locaux et en particulier du rôle des « bureaucraties de proximité » (*street level bureaucracy*). Cette nouvelle perspective a beaucoup emprunté à la sociologie des organisations qui permit un renouvellement des questions de recherche principales. C'était moins la mesure des écarts entre un projet de politique publique (ses intentions) et ses résultats objectivables qui était proposée, que des analyses sur les formes et les conditions de la concrétisation de ces programmes. Le regard évaluatif a ainsi été inversé. Délaissant une focalisation quasi-exclusive sur la quantification des résultats obtenus (efficacité), et encore plus sur celle de leur efficience (coûts/bénéfices des effets produits), ils furent centrés sur la caractérisation et la compréhension des processus qui sont la matière principale de l'action publique. Ils ont aussi développé un point de vue distancié, voire critique des décisions gouvernementales en mettant en évidence l'immensité des difficultés créées par : le flou des programmes, leurs incohérences (écarts entre objectifs et moyens alloués), mais aussi les différents modes d'appropriation et de réinterprétation des programmes effectués par leurs différents destinataires (les metteurs en oeuvre et les publics-cibles). Trois travaux célèbres incarnent cette perspective (Derthick 1972; Pressman, Wildavsky, 1973 ; Hardgrove 1975).

Plus récemment des travaux étasuniens poursuivent la réflexion sur les relations existant entre évaluation et système de pouvoir. Dans *The politics of policy evaluation*, M. Bovens et al (2006) montrent comment l'évaluation est dans la période contemporaine un enjeu entre un ensemble d'institutions qui revendiquent chacune des méthodes et des capacités d'analyse spécifiques pour pouvoir agir plus fortement sur le « policy making » : « *In the ideal world of the positivist social scientist, we stand to gain from this multiplicity : presumably it results in more facts getting on the table, and thus a more solid grasp of what happened and why. In the real world, multiple evaluations of the same policy tend to be non-cumulative and non complementary. Their methods and findings diverge widely, making it hard to reach a single authority or at least consensual judgement about the past and to draw clear-cut lessons from it* » (p.321). Estimant avec « *a slight exaggeration, paraphrasing Clausewitz, that policy evaluation is nothing but the continuation of politics by other means* » (p.321), ils analysent l'« *evaluation game* » institutionnel avec ses positions d'acteurs, leurs stratégies, leurs compétitions et

leurs “*stratagems of bureaucratic and budgetary politics*” (p.323). Par ailleurs, ils font le point sur la diversité des approches évaluatives et considèrent que les deux grands courants historiques que nous avons présentés plus haut structurent toujours la production intellectuelle dans ce domaine. Ils distinguent la « *rationalistic tradition* » (étude des performances) qui s’efforce de préserver l’évaluation des interférences politiciennes (p.325 s.) et l’« *argumentative tradition* » qui fait de l’évaluation une composante des débats publics et des affrontements d’intérêts (p.326).

France : un contre-pouvoir face à la technocratie budgétaire

En France la question des liens entre évaluation, contre-pouvoir et légitimité n’a été posée que beaucoup plus tardivement et dans des termes différents. La thèse de Vincent Spenlehauer «L’évaluation des politiques publiques : avatar de la planification» (Université de Grenoble, 1998) demeure le meilleur travail de recherche sur les transformations de la culture de la haute administration française à l’occasion de la planification (cf. encart n° 1). Il analyse aussi avec beaucoup de précisions les différentes conceptions du savoir à l’œuvre au sein du Commissariat général du plan, la tentative inaboutie d’importation des pratiques d’évaluation étasuniennes ainsi que les essais de rationalisation de l’action publique menés par le Ministère des finances avec la Rationalisation des Choix Budgétaires.

La période RCB (1965-1977) de pilotage des dépenses publiques basé sur des jeux d’indicateurs et d’indices est très peu relayée politiquement (Thoenig, 1971). Cette approche reste cantonnée au rôle d’outil de rationalisation du Ministère des Finances et de la Direction de la prévision. Ce type d’évaluation est pensé comme « une ressource de puissance technocratique ». Elle ne pourrait être exercée que par les « seuls acteurs qui disposent de l’expertise suffisante, qui échapperaient aux naïvetés du politique et à sa vénalité électoraliste » (Donzelot, Viveret, 1989, p.45). Les conclusions des études et leur possible traduction budgétaire deviennent une menace pour les autres administrations qui développent différentes formes de résistances tant technique que politique. Les Ministres soutiennent de fait les freins à ces mesures qu’ils perçoivent comme mettant en cause leurs prérogatives (Spenlehauer, 1995, p. 26-44). Les actions revendiquant le label RCB sont suspendues progressivement à partir du milieu des années 70. Le

pouvoir d'arbitrage décisif du Ministère des finances et du budget n'en est pas réduit pour autant.

Encart n° 1 : Résumé de la thèse de Vincent Spenlehauer

Cette thèse retrace l'histoire du thème de l'évaluation des politiques publiques au sein de l'Etat central français. La période étudiée s'étend de 1946 à nos jours. L'objectif du travail est d'apporter des éléments de compréhension et d'explication à l'échec de la politique nationale de développement de l'évaluation lancée officiellement en 1990 à l'initiative du Premier Ministre Michel Rocard. Ce travail énonce et valide l'hypothèse descriptive suivante.

Le thème de l'évaluation, importé des Etats-Unis par les planificateurs centraux français dès les années 1960, a été façonné et reste fortement marqué par les diverses formes de mythification de la planification nationale « performées » par les acteurs de l'Etat central français. Le caractère aujourd'hui globalement négatif que prennent ces entreprises centrales de mythification est à l'origine de l'échec de la politique nationale d'évaluation évoquée ci-dessus.

Cette thèse fonde finalement l'hypothèse interprétative suivante. En matière de technologies françaises de rationalisation de la conduite des affaires publiques, l'influence américaine a été grande dès la fin de la seconde guerre mondiale. Cependant, cette spécificité politico-administrative française que constitue l'existence des grands corps de l'Etat a opéré un filtrage de plus en plus fin des techno-sciences gouvernementales américaines, au point qu'aujourd'hui les sciences sociales françaises contribuent peu à l'évolution des cadres rationnels de l'action publique.

(Source : http://tel.archivesouvertes.fr/docs/00/34/25/92/PDF/These_doctorat_Vincent_Spenlehauer_1998_l_evaluation_des_politiques_publices_avatar_de_la_planification_.pdf)

Un deuxième moment intervient dans un contexte politique très différent à partir du début des années 80. Une série d'initiatives attestent d'une attention nouvelle aux questions d'évaluation. Elles partagent toutes un souci de validation démocratique des programmes d'action publique et une volonté de réalisation de réformes modernisant l'action administrative. En décembre 1983 est organisé de façon conjointe par le Ministère de l'Economie et des Finances et par le Secrétariat d'Etat à la fonction publique un colloque qui fait un bilan des pratiques de l'évaluation en France et tente la définition d'une stratégie globale pour son développement. La première originalité de la dé-

marche réside dans le tandem formé par Robert Poinsard (responsable de la cellule RCB du Ministère de l'économie et des finances) et de l'économiste Jean Pierre Nioche (professeur à HEC). La seconde repose sur les matériaux utilisés à savoir près de 300 études réalisées depuis la fin des années 70 dans différents organismes publics parallèlement à la RCB. Leurs conclusions gardent aujourd'hui une grande pertinence :

- L'évaluation est faiblement identifiée comme une activité spécifique distincte du contrôle autant que de la recherche. Elle est dès lors peu visible.
- Elle souffre d'une dispersion des pratiques, d'une absence d'homogénéité des profils des évaluateurs et des conceptions même de l'évaluation. De là, découle très peu d'échanges, de circulation de modèles et de confrontation avec les activités de ce type à l'étranger.
- La qualité méthodologique des travaux est très inégale. Les outils sont souvent rudimentaires, les démarches peu formalisées, avec une absence d'expérimentations systématiques. Les données produites sont de qualité faible avec beaucoup de manques et l'expression de beaucoup de difficultés à les produire. Les évaluations sont beaucoup plus descriptives qu'interprétatives, peu comparatives et insuffisamment explicatives.
- La rédaction manque de rigueur et d'objectivité : « Il apparaît souvent que la présentation et la discussion des hypothèses, des instruments et finalement des résultats obtenus s'effacent devant l'affichage des faits et des arguments en faveur de la thèse retenue »
- Les rapports entre les commanditaires, les responsables de l'action évaluée et les évaluateurs ne sont pas clairs. La confusion des responsabilités est fréquente.
- L'évaluation est peu insérée dans l'action politique et administrative. Elle remplit surtout un rôle d'information et de production de connaissances : « Son rôle dans les processus de décision et dans le débat public est actuellement faible ».

Bernard Perret en 1999 conclut que « ce diagnostic fournit un bon point de départ pour analyser les évolutions survenues depuis 15 ans ». En dépit de certains progrès, nous verrons que les principaux problèmes soulevés par

Nioche et Poinard ont conservé toute leur actualité. L'évaluation demeure une activité faiblement identifiée, insuffisamment systématisée, et dont les résultats sont sous utilisés (Perret, 1999, p.3).

A la même époque, le Parlement reprend le modèle anglo-saxon des «*technologies assesmment*» en créant en 1984 l'Office parlementaire d'évaluation des choix scientifiques et techniques (OPCEST) (cf. 1.5). La même année, le Commissariat du Plan met en place un groupe de travail présidé par M. Michel Deleau (sous directeur à direction de la prévision au Ministère des finances). Le rapport qui est produit reste une synthèse de référence pour la réflexion menée par les administrations centrales sous le double angle de la recherche d'une meilleure efficacité et d'une plus grande légitimité démocratique des choix publics (Deleau, 1986). A la suite de ces travaux, le Commissariat du Plan entreprend des « études évaluatives sur les politiques sociales transversales ». Dans ce contexte, plusieurs Ministères renforcent ou mettent en place des activités évaluatives de leurs programmes (cf. 1.3). C'est en particulier le cas du Ministère de l'Équipement qui fait appel à des chercheurs pour contribuer à l'approfondissement des cadres théoriques et méthodologiques. Les actes d'un séminaire et d'un colloque réalisés en 1983 en témoignent (Poinard, Nioche, 1984).

En filigrane de cet ensemble de réflexions et de pratiques, la dimension participative de l'évaluation est évoquée. L'accord est général pour rompre avec les évaluations hiérarchiques *top down*, qui sont considérées comme trop proches des inspections classiques et des *audits*. Ce point de vue sur l'action publique est jugé unilatéral et se cantonnant aisément dans la mesure de conformité ou de coûts, sans aborder vraiment les rouages complexes, les bails et les obstacles à sa concrétisation. Régulièrement, il est fait référence à l'intérêt qu'il y aurait à associer les destinataires de l'action au jugement porté sur elle. Mais qui sont ces « destinataires » ? Les points de vue oscillent entre, d'un côté « les bénéficiaires » ou « les régulés » et, d'un autre côté, les « metteurs en œuvre » chargés d'opérationnaliser les programmes publics. En tous cas, la pratique évaluative est incitée à s'ouvrir à des non experts, praticiens et/ou population concernée.

Un point de vue plus tranché est alors développé par P. Viveret à l'occasion de la mission que le Premier Ministre M. Rocard lui confie en juillet 1988. L'enjeu du RMI, politique expérimentale révisable, est l'occasion d'une réflexion et de la formulation de propositions plus larges sur les procédures d'évaluation. Dans le rapport publié (Viveret, 1989) l'auteur écarte ce qu'il nomme les approches techniciennes et positivistes qui prétendent « dire le vrai » sur une politique. Selon lui, la bureaucratie et ses normes, et à plus forte raison la technocratie et ses critères managériaux, se sont substitués aux acteurs politiques dans l'appréciation et l'orientation concrète des politiques publiques. L'évaluation doit permettre à ces derniers de reprendre la main et elle devient ainsi un contre-pouvoir à la technocratie. Il retient une définition qu'il veut politique et qui fait de l'évaluation « une conduite destinée à porter un jugement sur la valeur d'une politique »¹¹. Pour cela, il valorise le rôle du Parlement qui à l'occasion de ses activités délibératives doit formuler des choix basés sur des valeurs fondamentales et à partir de là « former l'opinion » au lieu de simplement en « recueillir l'adhésion » (Donzelot, Viveret, 1989, p. 46). Il propose aussi une revalorisation des services du Premier Ministre pour en faire un véritable lieu de synthèse des demandes sociales et des informations sur l'action publique afin de « procéder à de véritables arbitrages politiques ». L'objectif est de rompre avec le monopole exercé par l'expertise budgétaire du Ministère des finances « seul véritable lieu de synthèse » et de rendre possible l'expression et la prise en compte « d'intérêts divergents ».

Quant à la forme de l'évaluation elle doit répondre à trois critères principaux (Viveret, 1989, 121) :

- une exigence méthodologique : l'évaluation doit se distinguer du contrôle dans son objet et sa finalité. Elle doit s'appuyer sur des recherches quantitatives et qualitatives sans s'y réduire. « Son ambition est de reconnaître et d'apprécier les effets spécifiques d'une politique »
- une exigence démocratique : aucune évaluation « ne peut prétendre porter un jugement exclusif sur une politique. Elle est un produit sou-

¹¹ Un entretien entre P. Viveret et le sociologue J. Donzelot explicite précisément les fondements de sa réflexion (Donzelot, Viveret, 1989)

mis au débat démocratique». C'est pourquoi elle doit être transparente autant dans ses méthodes, ses informations que dans ses critères de jugement.

- une exigence d'indépendance : l'instance d'évaluation doit être indépendante de l'instance gestionnaire. L'auto-évaluation est souhaitable et doit être développée au niveau national et local, mais « elle ne saurait tenir lieu d'évaluation externe indépendante ».

L'évaluation du RMI de façon concomitante à sa mise en place (dispositif prévu par la loi qui l'a créé¹²) correspondra pour l'essentiel aux recommandations de P. Viveret. Mais elle reste en France, l'exception qui confirme la règle d'une évaluation très massivement technocratique et souvent maîtrisée par les gestionnaires des programmes. La mise en œuvre du RMI a suscité un vaste ensemble de travaux aussi bien d'ordre budgétaire, économique, organisationnels et sociaux (dont des études régionales comparatives) ont été réalisées. Une des synthèses les plus complètes se trouve dans MIRE et Plan urbain : « le RMI à l'épreuve des faits : territoires, insertion et société » (SYROS 1991). *A contrario*, la politique de la ville développée à la même époque n'a pas du tout donné lieu au même ensemble de travaux et d'études d'évaluation coordonnés (cf. encart n°15).

1.2. Les questions de modernisation de l'action publique et de compréhension des administrations

La culture de l'implémentation comme essor de l'évaluation

Outre la recherche de validation démocratique des politiques publiques, l'amélioration de la connaissance des mécanismes d'action publique et de compréhension des administrations chargées de la conduire a constitué une source d'inspiration majeure pour le développement de l'évaluation.

Cette dynamique peut être abordée selon deux mouvements historiques successifs. Dans un premier temps, sous l'impulsion des *Policy analysis* et l'analyse séquentielle rendue célèbre par O. Jones, l'évaluation a été

¹² Art. 52, loi RMI, 1er dec 1988

appréhendée comme une séquence spécifique des politiques publiques (Jones, 1970). Harold Laswell a été le premier, en 1956, à prendre en compte et à analyser une politique ou un programme gouvernemental comme un processus, c'est-à-dire un ensemble de phénomènes organisés dans le temps et animés par un certain nombre de mécanismes propres (Laswell, 1956).

Il découpe le processus en sept étapes distinctes : la compréhension (*intelligence*) qui correspond à la phase d'accumulation et de circulation d'informations auprès des décideurs, la promotion d'options par les décideurs, la prescription d'une façon de faire par les décideurs, la contrainte (*invocation*) qui correspond à l'établissement de sanctions pour adopter les mesures prescrites, l'application de mesures par l'administration et les tribunaux, l'achèvement (*termination*) de la politique ; l'évaluation constitue la dernière étape clôturant ce cycle ou qui le plus souvent le relance. Si ce modèle, fortement centré sur la décision a été largement remis en cause à partir de la fin des années 70, il a cependant eu le mérite de mettre en évidence la complexité de l'action publique et de faire de l'évaluation une partie prenante du processus. L'avantage de cette analyse séquentielle est de mettre de l'ordre dans « la complexité des actions et décisions constituant une politique publique ». Elle permet d'identifier la politique publique comme un objet « ayant une logique propre, identifiable et susceptible d'être analysé ».

La remise en cause des *Policy analysis* et le développement de l'analyse des politiques publiques a constitué le second mouvement de rattachement de l'évaluation. Il a permis de mettre en exergue la crise du paradigme balistique de l'action publique, notion qui est définie comme le fait qu'une action publique vise une cible et définit une trajectoire déterministe pour atteindre cette cible, comme dans le cas d'un tir d'artillerie. Dans le cadre de cette approche « balistique », le gouvernement agit en fonction de réponses prévisibles de la société. La remise en cause de cette approche dominante est non seulement la conséquence du fait qu'une société complexe est plus imprévisible, mais elle est surtout l'effet de la prise en compte de l'importance de la mise en œuvre. Cette approche a pour objectif d'analyser les effets concrets des décisions publiques, d'affiner la connaissance des modalités d'application et surtout des résultats de celle-ci.

La mise en œuvre peut être définie comme la phase de l'action publique durant laquelle des actes et des effets sont générés à partir d'un cadre

normatif d'intentions, de textes, ou de discours. La mise en œuvre désigne « l'ensemble des processus qui, après la phase de programmation, visent la réalisation concrète des objectifs d'une politique publique » (Knoepfel, Larrue, Varone, 2001, p. 214-215).

A partir des années 70, des auteurs, d'abord américains, ont été amenés à reconsidérer la phase de mise en œuvre des politiques publiques. Les grands ouvrages qui ont marqué ce domaine se sont basés sur des approches empiriques de l'action publique au concret en prenant généralement pour diagnostic préalable les constats d'échecs des politiques publiques relevant des Etats-Providence.

Ces travaux se sont tout particulièrement attachés à rendre compte des écarts entre l'idéal-type weberien de la bureaucratie et l'observation empirique des structures administratives. Les résultats des analyses des différentes politiques sectorielles ont permis d'interpréter les situations d'échecs des politiques publiques selon trois perspectives : l'ineffectivité, l'inefficacité et l'inefficience¹³ (Lascoumes, Le Galès, 2012).

En France, le mouvement a été moins net car jusqu'au début des années 1960, l'administration n'est pensée qu'en termes institutionnels et juridiques. Elle donne lieu ensuite à deux pistes de recherche. D'une part, des approches économiques en termes d'analyse cout-efficacité se développent dans le sillage des recherches de réforme de l'Etat (Laufer, Paradeise, 1982) et incarné par la rationalisation des choix budgétaires (RCB) (Bezes, 2003). D'autre part, un courant de sociologie des organisations parvient, autour de Michel Crozier et des chercheurs du Centre de Sociologie des Organisations (CSO) et sur la base d'études de cas significatives, à importer et adapter au contexte français les questionnements sur la rationalité et l'efficacité de l'action publique. Toutefois, ces travaux se sont plus attachés à rendre compte du fonctionnement concret des administrations (Crozier, 1963) ou d'institutions régionales (Thoenig, 1973 ; Gremion, 1976) et surtout à mettre au jour l'importance des dysfonctionnements bureaucratiques qu'à inscrire leurs recherches dans le cadre des politiques publiques à proprement parler.

Rassemblé autour de l'équipe de Jean Leca à Grenoble, le CERAT (Centre d'Etudes et de Recherche sur l'Aménagement du Territoire) constitue un

¹³ sur la définition de ces notions cf. *infra* p. 55 in 2.1

second lieu de naissance de l'analyse des politiques en France. Partant d'études de terrain sur l'activité de planification notamment, les chercheurs de ce laboratoire, derrière Lucien Nizard, et autour de Bruno Jobert et Pierre Muller, soulignent à la fois l'incohérence relative de *L'Etat en action* (Jobert et Muller, 1987) et le rôle de l'Etat dans la structuration de systèmes de représentations et de normes globales, ce que Jobert et Muller baptisent les « référentiels » d'action publique.

Développement et paradoxes de l'évaluation compréhensive en France

Telle qu'elle s'est structurée, en France du moins, au début des années 1990 l'évaluation s'est appuyée sur une sociologie compréhensive des administrations et de l'action publique. Aux côtés des approches visant à renforcer la démocratie, cette démarche a nourri le développement de l'évaluation sur la modernisation de l'action publique. Le rapport Viveret issu de la circulaire Rocard pour « le renouveau du service public » (23 février 1989) peut être considéré comme le ferment de cette « culture de l'implémentation » dans les activités d'évaluation en France, attentives aux écarts qui sont introduits nécessairement entre les intentions, leur mise en pratique, et les effets sociaux qui en découlent.

La politique ambitieuse de réforme administrative de « Renouveau du service public » engage un projet de modernisation de l'action publique articulé autour de quatre objectifs : le développement des responsabilités, la rénovation des relations de travail, l'évaluation des politiques publiques et l'accueil et le service des usagers.

Dans ce contexte est créé, par le décret du 22 janvier 1990, le Comité interministériel de l'évaluation (CIME), le Conseil scientifique de l'évaluation (CSE) et le Fonds national de développement de l'évaluation (FNDE). Ce décret reflète plutôt une conception gestionnaire puisque, selon ce texte, « l'évaluation a pour objet de 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 assignés ». Les objectifs en eux-mêmes ne font pas l'objet de l'évaluation.

Le dispositif ainsi mis en place en 1990 matérialise physiquement le lien entre la « *policy reseach* » et l'évaluation. De nombreux chercheurs spécialisés dans l'analyse des politiques publiques tels que Jean-Claude

Thoenig, Eric Monnier, Jean Leca et Patrice Duran sont ainsi associés aux réflexions menées par ces instances. Outre le crédit scientifique qu'ils apportent au développement de l'évaluation comme discipline, ils participent également directement au pilotage de ces structures naissantes. Ainsi c'est Jean Leca qui est choisi en 1990 pour Présider le Comité Scientifique de l'Evaluation (CSE).

Au-delà du dispositif institutionnel CIME/CSE, les approches en termes de compréhension de l'action publique ont également connu une importante mise en pratique dans différents secteurs d'action publique (cf. 1.3). Ces évaluations réalisées au niveau technique des ministères (Education, Culture, Equipement...) ont mobilisé et continuent de mobiliser les évaluations (souvent appelées « études évaluatives ») à des fins de production de connaissances et d'amélioration de leur efficacité.

Deux colloques tenus en 1992 attestent de la dynamique intellectuelle suscitée par le Conseil scientifique de l'évaluation (CSE). A ces deux occasions, ont eu lieu des rencontres entre responsables de l'évaluation dans des institutions publiques et des chercheurs universitaires. Les principaux noms dont les travaux ont marqué la période sont présents :

- le 15 juin 1992 a lieu dans le prestigieux centre de Cerisy un colloque organisé par le CSE sur « La modernisation du service public ». Une note introductive décapante de Sylvie Trosa « L'évaluation des politiques publiques : la qualité française » présente un bilan de la réflexion menée à partir de l'expérience du CSE. L'évaluation est présentée comme un jeu à trois acteurs (le commanditaire, l'instance, les chargés d'évaluation) qui est pour être maîtrisé doit reposer sur des interactions continues. La note analyse aussi les multiples résistances de la haute fonction publique à cet exercice. Elle conclut ainsi : « Toutes ces considérations s'appuient sur une histoire assez courte, et ne valent pas évaluation ! Elles le deviendront pourtant si l'évaluation connaît le sort de nombre de réformes administratives, c'est à dire d'être mises en sommeil avant que d'avoir eu le temps de se développer »

- le 2 octobre 1992 un autre colloque est organisé à Chantilly, toujours à l'initiative du CSE en partenariat avec le « Club de l'évaluation »¹⁴. Les principaux responsables de l'évaluation dans les services d'inspection ou d'évaluation administrative, et dans les Cour des comptes nationale et régionale, l'Insee etc. débattent avec des universitaires (J. Leca – Sciences-po, F. Pavé - CSO, E. Monnier-CEOPS, P. Lascoumes-GAPP). Le directeur général de l'administration et de la fonction publique, B. Pêcheur, conclue les travaux.

Mais paradoxalement, le succès de cette approche de « *policy research* » va créer une certaine forme de distanciation par rapport aux conceptions de l'évaluation des politiques publiques telles qu'elles vont être institutionnalisées et diffusées à partir des années 1990. En effet, parce que le regard porté par les analystes des politiques publiques sur l'évaluation reste profondément imprégné de cet idéal d'une action publique dont la capacité à transformer le social, à « résoudre des problèmes » s'avère faible voire aléatoire les différentes écoles d'analyse des politiques publiques en France vont entretenir un rapport ambigu aux pratiques d'évaluation. Sans être forcément toujours critique sur l'évaluation telle qu'elle va être massivement pratiquée à partir des années 1990 (cf. conclusion), les analystes vont insister sur les limites de l'évaluation à la française avant de s'en détacher progressivement à partir des années 2000.

Plusieurs principaux points de tensions caractérisent notamment le rapport des chercheurs à la dimension praxéologique de l'évaluation comme conseil aux décideurs publics.

- Tout d'abord, le lien avec la décision et ce que Jean Leca appelle « le paradoxe de l'évaluation » reste en effet problématique pour bon nombre de chercheurs de cette époque : « *Paradoxe de l'évaluation pour le décideur : quand elle est possible pour le décideur, elle est inutile ; quand elle utile pour le décideur, elle est impossible à mener* » (Leca, 1993).

¹⁴ Créé en 1991 le Club de l'évaluation regroupe les responsables de l'évaluation dans les ministères pour des échanges réguliers sur les concepts et méthodes.

- Ensuite, comme le note P. Duran (2010), on peut considérer aussi qu'une partie de ces réticences à l'évaluation peuvent relever du fait qu'« évaluation et organisation sont pour une bonne part, des termes contradictoires. En effet, l'évaluation n'est pas concernée par la survie de l'organisation : pour elle l'organisation n'a de validité que pour autant qu'elle a une utilité sociale, c'est-à-dire qu'elle est capable d'apporter une réponse pertinente à un problème public. Mais peut-on obtenir le support de bureaucraties quand l'analyse d'une politique conduit à remettre en question leur existence même ? » (p.6). Wildavsky avait précocement relevait cette contradiction en soulignant de manière volontairement provocatrice « qu'il est difficile de convaincre des fonctionnaires de collecter des informations qui peuvent leur porter tort ! » (Wildavsky, 1972).
- De plus, le développement d'un marché de service aux mains des prestataires privés a sans doute également contribué à défaire ce lien entre la recherche publique et l'évaluation. Pour Pascale Laborier et Daniel Gaxie (2003) cette expertise a principalement été confiée à des cabinets de consultants privés qui « sont souvent placés sous la dépendance commerciale du donneur d'ordre et qui dans un souci de rapidité et de rentabilité se contentent d'études relativement sommaires » (p. 211).
- Une quatrième hypothèse peut également être formulée selon laquelle le passage progressif à partir des années 2000 à des évaluations à visée démocratiques et de plus en plus « *bottom up* » a reconfiguré les rapports des autorités publiques à l'évaluation. En effet, les approches fondatrices basées sur les *policy analysis* et les sciences administratives ont été concurrencées par des courants de recherche plutôt axés sur l'étude des mécanismes de concertation et de démocratie participative (Cf. 1.4)¹⁵.

¹⁵ On peut citer par exemple J.-C. Thoenig peu convaincu par les approches démocratiques : « *L'évaluation ne mérite pas l'excès d'honneur qu'on lui réserve en disant que c'est la panacée qui permet enfin d'être démocratique ; non c'est une opération qui peut concourir à un meilleur fonctionnement des appareils publics comme à une meilleure appropriation de l'action publique par les différents partenaires de l'action publique (dont les citoyens, les publics cibles)* » (Congrès SFE, Rennes 2000)

1.3. Le renforcement des secteurs

En France, l'essentiel des activités d'évaluation a été développé au sein de Ministères et secondairement d'Agences (surtout agences de l'eau à partir de la fin des années 60). Nous donnerons quatre exemples qui permettent de saisir des contextes et des objectifs différents. Les trois premiers (Culture, Défense et Education nationale) sont les plus anciens et débutent dans les années 60. Le troisième (Equipement) a joué un rôle de modèle en France et comme pour la plupart des autres ministères il a développé des activités de production de connaissance à partir des années 70.

L'évaluation au service de la construction d'un secteur d'action publique : le Ministère de la culture

C'est l'exemple pionnier en France montrant l'articulation possible, même avec des moments de tension, entre l'action politique et la production de connaissance. Il est intéressant d'observer que cet exemple n'est à peu près jamais repris dans les travaux classiques sur l'évaluation car c'est la notion « d'études » introduite initialement qui a toujours été maintenue. La stabilité de ce choix terminologique s'explique en grande partie par la continuité des soutiens politiques (au-delà des très nombreux changements de ministres et de majorité) dont a bénéficié cette activité. Dans le domaine culturel, l'évaluation n'a jamais été ressentie comme une menace. Au contraire, elle a contribué à édifier et à entretenir la légitimité de ce secteur fragile de l'action publique.

« L'introduction du chiffre dans le domaine culturel est la condition nécessaire de l'introduction de l'objectivité et de la rationalité dans les décisions culturelles. Or sans cette objectivité et cette rationalité, le développement culturel ne saurait avoir la place qui lui revient dans le développement économique et social. La vie culturelle est si chargée de valeurs implicites, l'action culturelle si dépourvue d'objectifs clairs et cohérents, que seul le chiffre permettra de décrire, comparer, expliquer et prévoir », (Girard, 1971, p. 2-3.).

Initialement, c'est très tôt dans le cadre de l'élaboration du IV^o plan (1962-1965) que fut pensé un premier bilan de l'action culturelle menée par le Ministère créé en 1959. Un groupe présidé par Augustin Girard¹⁶ est chargé de ce travail. Cette démarche conduit à la mise en place en 1963 d'une « Cellule d'études et de recherches » dans ce ministère qui deviendra en 1968 le « Service d'études et de recherches » (SER), puis en 1986 le « Département des études et de la prospective » (DEP) et, enfin, en 2004 « Département des études, de la prospective et des statistiques » (DEPS). La production de ces services est très diversifiée et contribue à la consolidation de la légitimité de ce domaine d'action publique où l'intervention de l'Etat a longtemps semblé contraire à la liberté exigée par la production artistique. V. Dubois a bien montré les difficultés posées par la définition des frontières de ce qui relève (ou pas) des « affaires culturelles » (Dubois, 1999). Les études plus pragmatiques élaborées par le SER (sous ses divers intitulés) se sont attachées à démontrer l'apport de la culture à la richesse nationale (Ex : Les comptes culturels de la Nation, 1971) et à dresser le bilan synthétique des programmes et innovations accomplis durant certaines périodes décisives « Les affaires culturelles au temps de André Malraux 1959-1969 » réalisé en 1989, ou « Les affaires culturelles au temps de Jacques Duhamel 1971-1973 » réalisé en 1993. Des enquêtes ont aussi été menées sur l'action culturelle des collectivités locales à partir de 1993 (Poirrier – Dubois, 2002). Le terme d'évaluation est peu employé dans ces travaux. Rétrospectivement, c'est pourtant dans ce registre d'analyse là que ces enquêtes ont été conçues et réalisées dans un « grand écart entre la recherche et l'administration » afin de mettre « les études au service de la décision politique » (Pfioger, 2010). La production de connaissance a été voulue et entretenue dans le but d'institutionnaliser la politique culturelle, c'est-à-dire d'en montrer les frontières et leurs contenus tout en les légitimant par la démonstration de ses différentes formes d'utilité sociale (du renforcement des images symboliques à l'éducation et la mobilisation des citoyens en passant par la justification des métiers de la culture).

¹⁶ Né en 1926, professeur agrégé d'anglais à l'Ecole Alsacienne en contact avec Bernard Anthonioz membre du cabinet d'André Malraux, il sera intégré dans les groupes de travail organisé par le Commissariat général du Plan. Avant d'être intégré dans le Ministère.

Prospective et évaluation pour l'avenir de la Défense

Un autre exemple d'action pionnière dans la même période est la création en 1965 du « Centre de prospective et d'évaluation » (CPE) du ministère de la défense (Sfez, 1970, p. 300). C'est le premier organisme institutionnel français exclusivement consacré au « futur ». Ses missions sont : la prospective stratégique, l'évaluation et la sélection des orientations des programmes d'armement à long terme, la définition de la doctrine française en matière de dissuasion nucléaire. En 1982, le GROUPES succède au C.P.E. Il reprend les mêmes attributions en intégrant une fonction de court terme, en particulier par la réalisation de notes de synthèse sur des sujets d'actualité. Il réalise aussi des études à caractère opérationnel. Son rôle est plus politique que le C.P.E. et il est placé sous l'autorité du ministre. En 1987, le service devient la Délégation aux études générales (D.E.G.). Outre les compétences précédentes, il contribue à définir la politique d'armement nationale et assure le suivi des négociations sur ce sujet. Il joue un rôle dans la directive opérationnelle de la recherche. Le service est alors rattaché au cabinet du ministre.

Enfin, en 1992, le D.E.G. devient la Délégation à l'action stratégique (D.A.S). La DAS anime les réflexions prospectives du ministère sur l'anticipation des menaces, la typologie des conflits potentiellement militaires et contribue à l'aide à la décision stratégique. Elle réalise des analyses politiques contribuant à l'appréciation de situations et à l'anticipation afin de déterminer les positions françaises sur les dossiers intéressant la défense. Elle participe à la définition de la politique en matière de désarmement, notamment nucléaire, et de la lutte contre la prolifération et participe aux négociations internationales sur ces matières. Enfin, elle contribue au maintien d'une expertise extérieure sur les questions internationales et de défense. Elle soutient financièrement la réflexion stratégique indépendante, en particulier celle menée par les instituts de recherche et les organismes académiques français et étrangers.

Un exemple d'instabilité institutionnelle : le Ministère de l'Education nationale

Une relative ancienneté dans les pratiques évaluatives ne signifie pas pour autant la stabilisation des attentes à l'égard de l'évaluation, ni ne

garantit la viabilité à long terme des structures. Un exemple significatif est fourni par l'empilement des instances en charge de l'évaluation au ministère de l'éducation nationale et les tensions techniques et politiques qui ont accompagné cette prolifération. Ce cas montre typiquement la fragilité politique dont souffre souvent ce type d'activité (Derouet, 2006).

Un premier service statistique conçu avec l'Insee est mis en place en 1964 (Service des études informatiques et statistiques, SEIS). Il doit contribuer à la gestion de ce service public et à l'analyse des effets des réformes. Le SEIS participe avec le CEREQ¹⁷ aux travaux de planification. Il connaît un premier développement avec la réforme de 1975 qui introduit le principe du collègue unique. Le service devient Service de l'information et des études statistiques, (SIES) dirigé par Cl. Seibel). Des dispositifs sont alors mis en place afin avant d'évaluer « l'égalité des chances » des élèves en fonction de leurs origines sociales. Il met en place les premiers de suivis de panels d'élèves en 1977.

L'évaluation change d'ampleur à partir de 1981 (Agulhon, 2006). Tout d'abord est créé en 1984 un Comité national d'évaluation des établissements publics à caractère scientifique, culturel et professionnel (CNE) sous la forme d'une autorité administrative indépendante (encart n°3). Il a pour mission d'évaluer l'ensemble des universités, écoles et grands établissements relevant de la tutelle du ministre chargé de l'enseignement supérieur. Mais le ministère souhaite disposer d'une instance qui lui soit plus accessible.

Ensuite, après plusieurs changements de dénomination (plus que de mandats : SIGES en 1982, SPRESE en 1984 dirigé par J.P. Boisivon) l'ancien service d'études statistiques devient en 1987 la direction de l'évaluation et de la prospective (DEP), nom auquel on ajoute en 2007 dans le contexte de la RGPP le terme de « performance » (DEPP). Durant la période 70-2000 les grands enjeux de l'évaluation se complexifient, à l'évaluation des élèves¹⁸ s'ajoutent celle des établissements, puis celle des maîtres.

¹⁷ Créé en 1970 en partenariat par les ministères de l'éducation nationale, l'économie et le travail.

¹⁸ Intensifiée à partir de 2000 par la mise en place des procédures PISA (*Program of international studying assement*) promues par l'OCDE ... et dans lesquelles la France tardera à s'impliquer.

Enfin, les précédentes structures sont en compétition plus ou moins explicite selon les époques, avec les divers services d'inspection de ce ministère. Ainsi, l'évaluation des politiques d'éducation est organisée par la loi d'orientation de juillet 1989 qui assigne à cette activité un triple objectif : faire progresser le système éducatif, diffuser les pratiques innovantes et rendre compte à la Nation. Cette loi confie l'évaluation à l'Inspection générale¹⁹ qui doit procéder à des évaluations à tous les niveaux du système éducatif (départemental, académique, régional et national), en concertation avec les services administratifs compétents, et en s'appuyant notamment sur les productions statistiques de la DEP. Un décret de juillet 1990 étend cette mission d'évaluation aux inspecteurs territoriaux.

Le légo organisationnel est encore complexifié avec la création d'un Haut Comité d'Evaluation de l'Ecole (HCEE) en novembre 2000. Il est rattaché au même titre que le CNE et l'Inspection générale à l'administration centrale du ministère de l'Education nationale. Il a pour mission d'effectuer un travail d'évaluation de deuxième niveau du système éducatif à partir des travaux effectués par les différents acteurs en charge de l'évaluation que sont les corps d'inspection au sens large, la DEP, mais aussi les enseignants pour ce qui concerne l'évaluation des élèves. Le champ de compétence du HCEE couvre aussi bien les questions d'évaluation des acquis des élèves que celles touchant à l'évaluation des performances des établissements ou des pratiques éducatives. Il peut s'intéresser à tous les niveaux du système éducatif, y compris l'enseignement supérieur, voire la formation continue des adultes. Au vu de sa composition tripartite, il devait constituer une sorte de « comité des sages » sur l'école²⁰. Le HCEE a conduit des évaluations à la fois quantitatives et qualitatives sur des sujets très variés²¹. Mais, il va être

¹⁹ Par ce terme, nous désignons le corps d'inspection qui réunit les deux inspections générales que sont l'IGEN (inspection générale de l'Education nationale) et l'IGAEN, devenue en 1999 IGAENR (inspection générale de l'administration de l'Education nationale et de la recherche).

²⁰ Le HCEE est composé de trois groupes. Le premier réunit les membres proposés par des organismes représentatifs politiques (Assemblée nationale, Sénat) et professionnels (CFDT, MEDF, SGEN, l'UNEF/FAGE, etc.). Le second comprend des experts français et étrangers (des inspecteurs généraux, des directeurs de recherche, ou des professeurs d'université principalement). Le troisième est constitué par des représentants de l'Education nationale ayant voix consultative (chef de service de l'IGAENR, doyen de l'IGEN et directeur de la DEP). Le HCEE se réunit à la demande du Président de la République. Le chef de service de l'IGAENR et le doyen de l'IGEN assistent de droit à toutes les réunions.

²¹ Dix-neuf rapports furent publiés depuis sa création ce qui représente deux à six rapports par an en moyenne, sur des sujets aussi variés que la réception et l'utilisation en France des évaluations internationales des acquis des élèves, le

rapidement l'objet d'une valse-hésitation qui caractérise régulièrement le sort de telles structures dans beaucoup de secteurs de l'action publique et atteste du poids des interférences politiques.

Encart n° 2 : Le Comité national d'évaluation (CNE)

Il est créé en 1984 et a le statut d'une autorité administrative indépendante²². Il a pour mission d'évaluer l'ensemble des établissements publics à caractère scientifique, culturel et professionnel : universités, écoles et grands établissements relevant de la tutelle du ministre chargé de l'enseignement supérieur. Il peut également, soit à son initiative et avec l'accord du ministre chargé de l'enseignement supérieur, soit à la demande du ministre chargé de l'enseignement supérieur, procéder à l'évaluation d'autres établissements relevant de la tutelle de cette autorité ministérielle. Si le Comité l'estime nécessaire, il peut demander à exercer sa mission d'évaluation à l'égard d'établissements d'enseignement supérieur et de recherche relevant d'un autre département ministériel ; l'accord du ministre concerné est alors sollicité par le président du Comité national d'évaluation. Tout ministre peut également soumettre à l'évaluation du Comité les activités d'établissements d'enseignement supérieur relevant de sa tutelle.

Le Comité national d'évaluation examine et évalue les activités exercées par l'ensemble des établissements, et par chacun d'entre eux, dans les domaines correspondant aux missions du service public de l'enseignement supérieur : la formation initiale et continue ; la recherche scientifique et technologique ainsi que la valorisation de ses résultats ; la diffusion de la culture et l'information scientifique et technique ; la coopération internationale.

Dans l'exercice de cette mission, l'analyse du Comité national d'évaluation porte sur l'ensemble des actions et des moyens mis en œuvre par les établissements dans le cadre de leur politique scientifique et pédagogique.

Les analyses du Comité national d'évaluation sont consignées dans des rapports publics élaborés par établissement et par thème. Les rapports par établissement sont adressés au ministre chargé de la tutelle de ces établissements. Ils sont, en outre, adressés aux responsables de ces derniers. Les rapports par thème sont adressés au ministre chargé de l'enseignement supérieur et aux autres ministres concernés.

Les activités du Comité font l'objet d'un rapport adressé annuellement au Président de la République. En outre, le Comité national d'évaluation dresse tous les quatre ans un bilan de synthèse sur l'état de l'enseignement supérieur. Ce bilan est adressé au Président de la République. Le Comité national d'évaluation organise lui-même ses travaux : il fixe le programme de ses activités et détermine la méthodologie de ses évaluations.

(suite page suivante)

²² Source : Ministère de l'Éducation Nationale

Le ministre chargé de l'enseignement supérieur peut attirer son attention sur toute question appelant une évaluation en raison de son intérêt pour la politique nationale de l'enseignement supérieur et de la recherche ou sur toute mesure utile à la coordination des activités d'évaluation dépendant de son département.

En pratique, les activités du Comité national d'évaluation consistent en plusieurs types de travaux :

- des évaluations d'établissements : le Comité national d'évaluation a évalué, entre 1986 et 2000, la totalité des universités françaises ainsi qu'une trentaine d'écoles. Depuis 2000, les évaluations des établissements (universités, écoles, IUFM) sont organisées selon un calendrier harmonisé avec celui de la contractualisation du ministère de l'enseignement supérieur ; ainsi, les rapports d'évaluation peuvent être utilisés par les partenaires des négociations des contrats quadriennaux (présidence et direction d'établissement, directions du ministère).

- des évaluations de site, qui prennent en compte les interactions d'un ensemble d'établissements sur un territoire donné (région, aire métropolitaine, etc...)

- des analyses transversales concernant des activités spécifiques liées à l'enseignement supérieur

- des évaluations transversales et comparatives, portant sur une discipline (La géographie dans les universités françaises ; Les sciences de l'information et de la communication ; La formation des cadres de la chimie), ou un type de formation (L'odontologie ; Le 3ème cycle de médecine générale ; La formation des pharmaciens en France ; Les formations supérieures en mathématiques orientées vers les applications).

- des évaluations concernant des établissements relevant de la tutelle d'autres ministères (écoles d'agriculture).

En outre, le Comité national d'évaluation établit et publie périodiquement un rapport sur son activité et sur l'état de l'enseignement supérieur et de la recherche. Celui-ci est transmis au Conseil national de l'enseignement supérieur et de la recherche (CNESER).

À la date de novembre 2006, le Comité national d'évaluation a publié plus de 240 rapports d'évaluation.

La pratique du président du HCEE (Claude Thélot spécialiste reconnu de l'évaluation dans le domaine éducatif²³) qui accompagnait d'une conférence de presse la publication de chaque rapport a très tôt été critiquée en tant que signe d'une indépendance manifeste vis-à-vis des décideurs politiques. Institutionnellement, le HCEE avait été créé pour trancher le débat entre les tenants d'une évaluation externe et purement scientifique (position de C. Allègre à l'époque) et ceux qui défendaient une évaluation interne et participative (position de C. Thélot). Des observateurs ont vu dans cette création un moyen de compenser la perte de ses fonctions à la tête de la DEP (1997) car il était alors considéré comme une véritable « voix » de l'école sous un ministère hostile à toute évaluation interne. A la même époque la DEP est rebaptisée « direction de la programmation et du développement ». La création du HCEE correspondait aussi aux attentes de nombreux évaluateurs (les inspecteurs généraux surtout) qui souhaitaient que leurs rapports aient davantage d'audience interne et d'effets.

Pourtant au bout de quatre ans de fonctionnement, le HCEE est supprimé à son tour en avril 2005 qui le remplace par le Haut Conseil de l'Education. Celui-ci doit intégrer d'autres missions à côté des travaux d'étude et d'évaluation. Il doit aussi élaborer le cahier des charges de la formation dispensée dans les IUFM (Instituts Universitaires de Formation des Maîtres) et constituer le fameux « socle commun de connaissances et de compétences » de base qu'on est en droit d'attendre de tout élève à la sortie de la scolarité obligatoire. Il intègre pour cela l'ancien Conseil national des programmes. L'évaluation est donc intégrée dans un dispositif plus large.

Ces changements à répétition conduisent à s'interroger sur la cohérence à moyen terme de la politique d'évaluation en éducation. Qu'attend-on d'elle ? Qui en décide ? Que fait-on de la connaissance produite ? Les fluctuations observées sur ces questions montrent que, au mieux elles fragilisent la production de connaissance sur l'action publique, au pire elles la rendent impossible. Une comparaison avec l'Allemagne est intéressante sur ce point. Les résultats des enquêtes triennales de l'OCDE à partir des années 2000 ont suscité un véritable « choc PISA » dans ce pays au vu des résultats de la première décennie. Le système éducatif a été profondément modifié en

²³ Inspecteur général de l'Insee, il a dirigé la DEP de 1990 à 1997. Il est l'auteur de *L'évaluation du système éducatif. Coûts, fonctionnement, résultats*, Nathan, 1994.

mettant l'accent sur la formation des enseignants et leur accompagnement. En dix ans les résultats ont été tangibles. En France, les spécialistes n'observent ni « choc », ni même « effet » PISA. Comme cela s'observe souvent c'est la méthode elle-même qui est contestée et les résistances à l'évaluation, en particulier celle des élèves, s'accroissent²⁴. Pourtant les évaluations nationales donnent des conclusions concordantes à celles effectuées par l'Ocde. Au-delà de la complexité considérable du système de l'éducation nationale, tant des professionnels de l'enseignement que des spécialistes de l'évaluation s'interrogent sur la portée politique de la masse considérable des données et analyses produites depuis, au moins trente ans, sur le système de l'éducation nationale (Normand, Derouet, 2011).

Un exemple d'évaluation au service de la modernisation : le ministère de l'Équipement

Les actions du ministère de l'Équipement en matière d'évaluation font très souvent figure de référence en France. Pourtant, il n'est pas historiquement pionnier et ses actions ne se développent que dans les années 70. Cependant, plus que beaucoup d'autre, il a su orienter la production de connaissance sur ses activités (coût, gestion des personnels, programmes innovants) avec une politique de modernisation de ses services (induite en particulier par la décentralisation) et de mobilisation des personnels (redéfinition des services, formation, gestion des carrières). Un élément de contexte général contribue à donner au statut de l'évaluation dans ce secteur administratif une place particulière, il s'agit de l'histoire longue, voire très longue, des relations entre la production de connaissance et la politique urbaine (Van Damme, 2005).

Le point de départ se situe dans le courant des années avec une série de programmes d'études sur l'urbain et la transformation des villes commanditées par le Commissariat du Plan. L'étape initiale est la création d'une Mission de la recherche urbaine (dirigée par Michel Conan) en 1973

²⁴ Pourtant, à l'entrée du secondaire, sur la base des tests Pisa on estime à 32% les élèves en difficulté. La moyenne de l'Ocde est 18%. A cela s'ajoutent deux phénomènes : la baisse de la proportion des élèves classés de très bon niveau ; et l'alignement tendanciel des résultats de filles (qui étaient longtemps meilleures) sur ceux des garçons.

qui devient le service d'analyse et de prospective du ministère²⁵. En 1984, il devient le Plan urbain, dirigé par André Gruson. Enfin la fusion avec le « Plan construction » aboutit en 1998 à la création du PUCA (Plan urbanisme construction architecture). Durant ces trente années cette structure est devenue le véritable « poisson pilote » du ministère de l'Équipement. Les rapports de recherche qu'elle produit ou suscite (via une importante politique d'appel d'offres), ainsi que son souci de valorisation des résultats à travers une revue « Les annales urbaines » et l'organisation de nombreux colloques et séminaires tissent des liens profonds avec les intellectuels de science sociale (sociologues et économistes surtout). A dix ans d'intervalle, deux livres ont analysé cette forte interpénétration atypique des milieux scientifiques et administratifs (Amiot M., 1986 ; Lassave P., 1997). En 1979, M. Conan devenu chef du département des sciences humaines et sociales du Centre scientifique du bâtiment recrute Eric Monnier qui a été un des grands promoteurs de l'action évaluative très au-delà de ce qui est réalisé au ministère de l'Équipement et qui garde valeur de modèle (Monnier, 1987).

Successivement ces services vont aussi s'engager dans des actions combinant la recherche, l'expérimentation et l'évaluation, en France, mais aussi à l'étranger (activité très originale). Ainsi M. Conan fait réaliser une évaluation des « communes libre » suédoises qui ont mis en place des systèmes de gestion locale participative indépendants. Une autre étude porte sur l'évolution des services urbains new-yorkais à la suite de leur crise financière de 1975. Une restructuration radicale a été négociée avec les personnels qui a aussi permis de fixer des objectifs de qualité.

A partir des années 80, le Plan urbain s'engage dans une réflexion très large sur les activités d'évaluation. Un séminaire de méthode associant universitaires, représentants de bureau d'études et d'autres administrations est mis en place en 1988. Il débouche en 1990 sur le premier guide français de l'évaluation.

Enfin, les services du Ministère de l'Équipement vont tenir un rôle très important dans l'évaluation des actions expérimentales telles que les régies de quartier (Roubaix 1985-1989), le Revenu minimum d'insertion et la

²⁵ Ce service est en partie l'héritier du STCAU (Service technique central de l'aménagement et de l'urbanisme) qui avait pris des positions très politiques en 1968 et demandé la dissolution du corps des Ponts et chaussées ... Il fut dissout quelques mois plus tard (Spénlehauser, 1995, p. 31)

politique de la ville. A partir de 1996, sous le thème « L'Etat local », une série d'enquêtes sont menées sur l'évaluation des actions des services territoriaux.

Un autre exemple de création d'un service d'étude ministériel lié aux efforts de modernisation de l'action administrative des années 70 est celui du « Centre d'analyse et de prévision » (CAP) du ministère des Relations extérieures en 1973. Il s'inspire de l'organisme de prévision du département d'Etat étasunien le *Policy Planning Staff* (PPS) (encart n°3). Un tel service était au départ prévu au sein du Commissariat Général du Plan. Ce rattachement échoue et la structure est finalement créée au sein du ministère des Relations extérieures, mais il est mis en place par deux transfuges du CGP (Spencehauer, 1998, p. 155-156). Les orientations du CAP évoluent dans le temps, en particulier en fonction des commandes ministérielles (Cohen, 1982). La mission prospective initiale est rapidement concurrencée par les activités d'aide à la décision censées rendre le service plus opérationnel et renforcer ses liens avec les autres services. Dans une première période (1973-1981), le CAP tient un rôle de contre-expertise en utilisant des sources extérieures au ministère. Il participe à l'analyse des diverses crises, en particulier pétrolières et à la détermination des orientations de politique étrangère qui en découlent. A partir de 1981, les demandes ministérielles se font plus globales et portent sur de grands axes stratégiques dans les relations internationales (espace social européen, dénucléarisation du Moyen Orient, politique africaine, etc.). En 2009 la dimension d'analyse prospective est affirmée par un changement de nom, le service devenant le « Centre d'analyse et de prospective » (CAP).

La dimension évaluative des activités du service est toujours restée minoritaire. De plus, les liens de proximité initiaux avec le Ministre et son cabinet sont toujours restés très forts. D'où un certain nombre de crise en particulier pour le choix des directeurs successifs (2009, 2012).

Encart n° 3 : Le « Policy Planning Staff » aux Etats-Unis

The Mission and Purpose of the Policy Planning Staff

Created in 1947 by George Kennan at the request of Secretary of State George C. Marshall, the Policy Planning Staff (S/P) serves as a source of independent policy analysis and advice for the Secretary of State. The Policy Planning Staff's mission is to take a longer term, strategic view of global trends and frame recommendations for the Secretary of State to advance U.S. interests and American values.

In his memoirs Present at the Creation, former Secretary of State Dean Acheson characterized the role of Policy Planning: "To anticipate the emerging form of things to come, to reappraise policies which had acquired their own momentum and went on after the reasons for them had ceased, and to stimulate and, when necessary, to devise basic policies crucial to the conduct of our foreign affairs."

For today's Policy Planning Staff, fulfilling this same, core mission for the Secretary requires striking a fine balance between engagement in the day-to-day requirements of diplomacy and development of long term, strategic plans. Broadly speaking, the daily work of the Policy Planning Staff may be divided into six areas:

Analysis

Policy Planning serves as an internal think tank for the Department of State - undertaking broad analytical studies of regional and functional issues, identifying gaps in policy, and initiating policy planning and formulation to fill these gaps. Policy Planning also serves as an institutionalized "second opinion" on policy matters - providing recommendations and alternative courses of action to the Secretary of State.

Special Projects

Policy Planning assumes special projects or takes the lead on certain issues as tasked by the Secretary of State. Recent examples include Policy Planning's work on assembling the international coalition against terrorism, coordinating the reconstruction of Afghanistan through February 2002, and implementing the Good Friday agreement in Northern Ireland.

Policy Coordination

Policy Planning engages functional and regional bureaus within the State Department and relevant government agencies to ensure coordination and integration of policy with longer-term objectives.

Policy Articulation

The speechwriters for the Secretary of State are members of the Policy Planning Staff and work together with the whole Staff and all bureaus to draft the Secretary's speeches, public remarks, testimony before Congress, and contributions to print media.

(suite page suivante)

Liaison

Policy Planning acts as a liaison with nongovernmental organizations, the academic community, think tanks, and others to exchange expert views on matters relevant to U.S. policy and to ensure that broad public opinion informs the policy formulation process.

Planning Talks

Policy Planning holds a series of dialogues -- known as planning talks -- with counterparts from other countries, including our key European allies, Japan, South Korea, Australia, China, and Russia. These talks provide an opportunity to discuss broad strategic issues that go beyond crisis management or the day-to-day concerns of diplomacy.

En France, l'essentiel des activités concrètes d'évaluation est réalisé au niveau institutionnel des ministères. Ils ont été, le Commissariat général du plan mis à part, les principaux lieux d'expérimentation et de mise en œuvre concrète de ce type de production de connaissance. Cependant, comme l'atteste les trois exemples présentés, il existe entre eux des disparités considérables. Globalement, le suivi des politiques en cours et à plus forte raison l'impact des évaluations sur les décisions prises restent faibles. Pour comprendre cela il suffit d'examiner la situation des départements ministériels considérés comme les plus avancés dans le domaine. Ceux qui sont reconnus comme les plus constants dans le recours aux démarches évaluatives et les mieux expérimentés méthodologiquement sont les Ministères de l'Équipement, de l'Éducation nationale et celui du Travail et des affaires sociales. Cette continuité s'explique surtout par la pression que ces secteurs subissent de la part des usagers et des demandes sociales. Ils n'ont pas pour autant résolu la question récurrente de la production d'une information cohérente continue sur les actions et programmes publics dont ils ont la charge. Un rapport parlementaire de 2004 (J. Bourdin, P. André, J.-P. Placade X., 2004) prend ainsi l'exemple de la politique de l'emploi, ensemble de programmes stratégiques s'il en est. Les analyses faites montrent le niveau pléthorique des dispositifs existants et l'incapacité des gouvernements successifs, non seulement à les homogénéiser, mais même à en comprendre l'enchevêtrement malgré des systèmes d'observation très nombreux mais qui sont restés très peu coordonnés. Depuis ces dernières décennies les mesures ont été empilées, juxtaposées sans évaluation préalable et l'avis des parlementaires rejoint le diagnostic de plusieurs spécialistes. Tous s'accordent aujourd'hui pour constater avec l'économiste

B. Gazier que l'on est « arrivé à un paradoxe : plus vous faites des politiques de l'emploi, moins vous pouvez les évaluer » (Gazier, 2003)²⁶.

Le rapport parlementaire conclut plus largement par une critique de l'éparpillement observé des activités d'évaluation dans les institutions publiques. Ils relèvent que si l'essentiel de la production de connaissance sur l'action publique est produit dans les ministères, celle-ci reste de qualité incertaine et peu cumulable. Ils expriment trois critiques principales.

Tout d'abord, ils « regrettent » que ce processus « se soit déroulé sans qu'une stratégie cohérente ne vienne l'inspirer, en dépit des travaux consacrés à l'évaluation par le Comité interministériel à la réforme de l'Etat ». Les quelques exemples donnés ci-dessus indiquent à quel point chaque service ministériel s'est développé en fonction de logiques strictement internes et sans aucune référence à des modèles extérieurs validés. Si on trouve des références à des modèles étrangers, elles restent très vagues, souvent rhétoriques et l'appropriation a été accompagnée de réorientations massives tant dans l'esprit que dans les moyens. De plus aucun modèle national n'a été élaboré. Chacun a œuvré selon les compétences internes disponibles et le classement de ces différents services existants oscille entre les structures de conseil au ministre à la prospective, en passant par la production de données internes, l'auto-évaluation et l'action de changement par apprentissage.

Ensuite, les parlementaires observent que « l'apport à l'évaluation des politiques (est) resté très limité ». Aucune capitalisation de méthodes, échange de savoir faire ou réflexion méthodologique transversale n'est observable dans la période relativement longue allant du milieu des années 60 à aujourd'hui. Tout se passe selon un précepte digne de Monsieur Jourdain qui n'apprend que de lui-même. En près de cinquante ans de pratique évaluative, il n'y a pas vraiment eu d'accumulation de connaissances. Les parlementaires considèrent que l'essentiel de la production de connaissance est du type « auto-évaluation ... De telles réalisations, qui n'offrent que très peu de garanties de qualité pour les

²⁶ Et dans le même sens voir Ph. Garraud (2000), *Le chômage et l'action publique, le bricolage institutionnalisé*, L'Harmattan.

commanditaires (quasi-exclusivement les ministres) ne peuvent, par nature, contribuer à la promotion de l'évaluation des politiques publiques »

Enfin, les parlementaires considèrent que le défaut majeur des activités évaluatives françaises est de rester pour l'essentiel sous l'emprise du pouvoir exécutif. Et cela autant pour le choix des programmes ou enjeux à évaluer que pour la conduite même de ces démarches : « Ces derniers constats (les deux points précédents) mis en regard de la situation de quasi monopole de l'expertise par l'exécutif posent la question des moyens d'optimiser les ressources attribuées à celles-ci ». Ils présentent alors leur proposition centrale, la création d'une Haute autorité de l'évaluation », proposition que nous analyserons dans la deuxième partie (cf. 2.4).

1.4. Les approches démocratiques : de l'évaluation pluraliste à l'évaluation émancipatrice

Parallèlement aux approches en termes de production de connaissances et de compréhension, la notion d'évaluation pluraliste est introduite en France par le rapport Viveret de 1989. Selon cette notion, la finalité de l'évaluation n'est pas seulement économique et financière, mais aussi politique. Fondant sa présentation sur des conceptions développées jusque-là aux Etats-Unis, en particulier par Guba et Lincoln (1989) (encart n°4), le rapport commandé à Patrick Viveret développe l'idée que l'évaluation doit être démocratique en organisant un débat social ouvert et transparent sur les choix de politiques publiques « loin de toutes tentations de la voir confis[quée] par l'Etat et les grands corps qui la considèrent comme une « ressource de pouvoir » (Spencehauer, 1998).

Plusieurs principes organisationnels découlent de cette conception démocratique et pluraliste. En premier lieu, le principe d'une association des différents acteurs concernés par l'évaluation, énoncé en particulier par le Conseil scientifique de l'évaluation qui insiste sur la confrontation des points de vue. L'évaluation répond également à une volonté d'approfondissement de la vie démocratique : l'Etat doit devenir responsable de ses actions devant les citoyens.

Encart n° 4 : Les quatre générations d'évaluation selon Guba et Lincoln (1989)

- 1^{ère} génération : De la fin du 19^{ème} siècle aux années 1920 : les évaluations se limitent à mesurer les effets de sanctions ;
- 2^{ème} génération : De 1920 à 1957 : les évaluations doivent également décrire les forces et les faiblesses du dispositif public ;
- 3^{ème} génération : De 1957 à la fin des années 1980 : l'évaluation doit également apporter un jugement de valeur sur le programme ;
- 4^{ème} génération : Depuis la fin des années 1980 : l'évaluation intègre la participation des acteurs.

L'exercice de cette responsabilité suppose de ce fait des instances d'évaluation indépendantes de l'État. Il s'agit dans ce cadre de porter des jugements de valeur sur les politiques publiques, y compris sur leurs objectifs : c'est là une différence forte avec l'évaluation gestionnaire ; l'évaluation peut être amenée à redéfinir les finalités de l'action publique. Celles-ci peuvent être redéfinies en fonction du déroulement du processus de mise en œuvre. Au cœur de ce mode d'évaluation se trouve donc l'idée qu'elle doit être un support au débat public (permettant la détermination des objectifs). En l'espèce, les expériences menées dès les années 80 en Suisse, et plus précisément au sein du canton de Genève, font office de référence (encart n°5). La Commission externe d'évaluation des politiques publiques (CEPP) installée à Genève réunit l'ensemble des forces de la société civile, le plus souvent non professionnelles de l'évaluation. Disposant d'un pouvoir d'auto-saisine, son rôle explicite consiste à « tendre un miroir à l'administration sur ses activités » (Delley, 2006, p. 46-49).

Ce mode d'évaluation correspond en troisième lieu à une volonté d'ouverture de l'administration vers le public, par une prise en compte du point de vue des ressortissants d'une politique publique. Les « *stakeholder evaluations* » (Chemlinsky, 1993) essaient ainsi d'examiner les perspectives de tous ceux qui sont concernés par une politique et par une évaluation. Sous cette perspective l'évaluation est considérée comme indispensable pour expliquer de façon persuasive et crédible la logique et les effets des décisions publiques. Dans une perspective pluraliste, ce type d'évaluation peut être

aussi conçu comme un lieu de négociation itératif et interactif. Cette évaluation est itérative car elle doit permettre de corriger l'action publique dans le cadre d'un processus d'apprentissage. Il y a ici en quelque sorte une circularité de l'action publique puisque l'évaluation, fondée sur la mise en œuvre, alimente aussi la décision publique. Elle est également interactive puisque plusieurs évaluateurs, voire plusieurs modes d'évaluation sont confrontés. Le processus est collectif et pluraliste.

Encart n° 5 : Les pratiques d'évaluation du canton de Genève

L'approche sans doute la plus poussée d'évaluation démocratique a été expérimentée en Suisse, plus précisément au sein du « micro-cas » que constitue le Canton de Genève. En 1985, le Canton de Genève a institutionnalisé l'évaluation de ses lois en créant en 1985 un organe permanent et extérieur à l'administration : la CEPP (Commission externe d'évaluation des politiques publiques). Ses membres, au nombre de 16, sont choisis en dehors de l'administration et de toute autorité politique du canton et proviennent des milieux économiques, académiques, médiatiques ou judiciaires (anciens députés ou magistrats). Cette instance est également pourvue d'un droit d'auto-saisine et qu'en Selon l'universitaire Jean-Daniel Delley qui en a été le Président pendant 8 ans, la CEPP joue le rôle d'une « commission de milice » dans la mesure où elle s'appuie sur toutes les forces de la société civile même si ces dernières ne sont pas des professionnels de l'évaluation (Lascoumes, Varone, Pons, 2006). A travers ces évaluations, dont la durée maximale est limitée à 1 an, le rôle de la CEPP consiste à « tendre un miroir » à l'administration sur ses activités. Le choix des thèmes d'évaluation ne doit être ni trop partisans ni trop lié à l'actualité mais doit être utile et proportionné aux enjeux à traiter. Plusieurs critères ont ainsi été mis en place par la CEPP : l'enjeu financier, l'exemplarité (par exemple l'évaluation des mécanismes de subventionnement des écoles de musique), le fort degré de conflictualité ou au contraire le grand consensus (Jacob, Varone, Genard, 2007, p 113-122)

En France, l'architecture d'ensemble du dispositif d'évaluation mis en place dans les années 90 au niveau étatique découle de cette approche. Si le CSE placé aux côtés du CIME et chargé d'assurer l'interface avec les acteurs de la société démocratique en est l'exemple le mieux connu, le dispositif prévoyait également un certain nombre de mesures placées dans le champ démocratique. Ainsi par exemple, le Conseil économique et social ou le Médiateur de la République étaient identifiées comme instances pouvant

faire l'objet de saisine pour des évaluations en provenance de la société civile

Dans le bilan qu'il dressait en 2008 des débats d'origine Patrick Viveret soulignait la force de cette impulsion démocratique initiale : « ce qu'a porté Michel Rocard, et c'est pour cela qu'il y a eu une énergie fondatrice originelle, c'est que l'évaluation est un acte démocratique majeur et fait partie d'un projet politique, au sens le plus fort et le plus noble du terme » (SFE, 2010).

Sur le plan organisationnel, cette évaluation démocratique n'a pas réussi à s'imposer. Cela provient d'abord des réticences voire des blocages des acteurs institutionnels comme le Parlement. Le poids et la surdétermination de l'exécutif inhérents à la Cinquième République ainsi que le développement d'une culture « sectorielle » de l'évaluation dans les administrations et les ministères ont également freiné la concrétisation du dispositif démocratique et transparent souhaité par le rapport Viveret.

« Le résultat c'est que comme on ne pouvait pas proposer officiellement quelque chose que le Parlement n'était pas prêt à accepter, eh bien, l'ensemble du dispositif que j'avais proposé à Michel Rocard s'est trouvé déporté, en quelque sorte malgré nous – je peux en témoigner du fait que Michel Rocard en était le premier meurtri – du côté de l'exécutif » (Viveret, 2008)

Dans les faits, cette vision de l'évaluation pluraliste a réussi s'imposé en France. Comme le souligne Jean-Claude Barbier, le CSE, bien qu'il n'ait jamais réuni les ressources adéquates pour asseoir une influence prépondérante, n'en constitua pas moins un référent majeur et réussit à produire cette doctrine de l'évaluation pluraliste (Barbier, 2004). Le CSE privilégia en effet souvent le choix de l'installation d'une instance pluraliste, représentant les divers points de vue qu'il qualifiait de « légitimes » à propos d'une politique, pour lui, la meilleure *stratégie cognitive*. 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, Setbon, 1996, p. 10-11). Cette idée est reprise par B. Perret qui plaide pour la diffusion des évaluations auprès du grand public à travers le passage de

« l'évaluation des politiques publiques à l'évaluation publique des politiques » (Bernard Perret, 2008). La Société française de l'évaluation a également, après de longs débats en son sein, repris cette préférence pour l'évaluation pluraliste.

Dans les pratiques en revanche, les expériences d'évaluation démocratique ont connu une vitalité indéniable. Après le passage progressif des évaluations managériales, faites pour et par la seule administration, les études et travaux d'évaluation développées depuis la fin des années 1990 ont eu pour souci constant d'associer des tiers hors de l'administration. Cette nouvelle vague d'évaluation se caractérise par la mise en œuvre de démarches plus ambitieuses en termes d'implication de la société civile qu'il s'agisse des évaluations participatives et depuis quelques années des expérimentations en faveur d'évaluations « émancipatrices » (*empowerment evaluations*).

Tout d'abord, le lien entre évaluation et participation a donné lieu, ces dernières années, à une production littéraire et scientifique importante. Parallèlement à la montée en charge des recherches sur la démocratie participative les réflexions sur la place de la participation des bénéficiaires, usagers ou citoyens dans les processus évaluatifs, si bien que l'on a assisté à une sorte de fertilisation croisée entre ces deux champs de recherches²⁷. Plusieurs ouvrages ou articles ont ainsi tenté de décrypter les relations entre « la participation » et l'« évaluation », sous trois angles principaux ayant permis d'étudier la contribution de la diffusion des résultats d'évaluation au débat public (Patton, 1996 ; Weiss, 1998), d'analyser l'apport de la participation des citoyens à des missions d'évaluation (Barron, Monnier, 2003) ou de théoriser des dispositifs permettant d'évaluer les politiques ou programmes en faveur de la participation des citoyens (Faure, 2004).

Ces recherches ont pour point commun de contribuer à affirmer la dimension éminemment démocratique des évaluations, reléguant au second plan la finalité managériale. Elles portent une attention particulière à la volonté de démocratiser et de rendre plus transparente la vie publique, ce qui implique de donner un accès plus libre aux dispositifs d'évaluation. De plus, les

²⁷ Pour n'aborder ici que les finalités des deux exercices, on pourra par exemple, s'étonner de la similarité des deux exercices qui se rejoignent dans leur rapport, théorique comme pratique, à l'action publique.

évaluations participatives ont été préconisées en réponse au constat de faible utilisation directe et opérationnelle des évaluations managériales : elles apportent simplement, en règle générale, une meilleure compréhension du fonctionnement du programme, et une contribution aux réflexions sur les actions publiques et sur l'évaluation. L'un des effets attendus de ces évaluations participatives est donc de permettre aux participants de mieux assimiler les résultats et analyses de l'évaluation pour accroître les chances de leur utilisation directe. Concrètement, si les expériences d'évaluation participatives se sont développées de façon naturelle d'abord dans le champ des politiques sociales (aménagement, développement social urbain...), les politiques environnementales, sous l'effet de l'impératif de développement durable ont constitué également un terrain d'application privilégié pour ces approches. La politique de l'eau associant des usagers, la mise en place des politiques climatiques (*climate change policies*) et surtout les évaluations de démarches de développement durable des territoires (Agenda 21, SRADDT, SCoT, Plans Climat-Energie Territoriaux...) en sont les exemples les plus connus.

Dans le sillage de ces démarches participatives, un courant encore plus récent d'évaluations à visée démocratique a émergé en Amérique du Nord, sans que la pratique soit encore très répandue. Ces approches, que G. Baron et E. Monnier qualifient de 5^{ème} génération d'évaluation en complétant la terminologie de Guba et Lincoln, vont plus loin et proposent que la société civile « prenne le pouvoir », autrement dit décide du contenu de l'évaluation, pilote elle-même les travaux, et produise un jugement sur les actions publiques (Baron, Monnier, 2003). Partant du constat que l'information et la connaissance sont un pouvoir trop souvent réservé aux décideurs et aux principaux acteurs d'un programme, ces évaluations considèrent indispensable de transférer ce pouvoir notamment aux groupes les plus démunis dans l'optique de les mettre en capacité ou de les émanciper (Fetterman, Kaftarian, Wandersman, 1995).

Cette approche, employée pour l'évaluation de programmes sociaux, éducatifs ou de renouvellement urbain aux Etats-Unis et au Canada ou de projets d'ONG dans les pays en voie de développement, vise certes à améliorer une politique publique, mais aussi à aider les groupes sociaux concernés à créer leur propre savoir, leur propre compréhension de leur situation et des moyens nécessaires pour la modifier. Pour que chacun des groupes participants soit en capacité d'accéder à la connaissance qui sera

produite au cours du processus évaluatif, les théories fondant ces approches prévoient une phase préalable conséquente de formation à l'évaluation, mais aussi à la prise de parole et à l'analyse. Elles s'appuient sur des méthodologies majoritairement qualitatives et davantage sur des dessins, des photos et des schémas que sur des écrits pour que même les groupes les plus démunis puissent assimiler les informations.

Paradoxalement, un bon exemple de cette approche nous est donné par une évaluation réalisée en France, dès 1987 : l'évaluation d'opérations de Développement Social Urbain. Probablement en raison de son caractère trop novateur cette expérience n'a malheureusement pas convaincu l'administration centrale de l'époque qui s'est probablement sentie dépossédée d'une partie de ses responsabilités (cf. encart n° 15)

Encart n° 6 : Les évaluations émancipatrices, une 5ème génération d'évaluations

Génération	Type d'évaluation	Qui décide du contenu de l'évaluation ?	Qui pilote les travaux d'évaluation ?	Qui produit le jugement final ?
1ère à 3ème	Managériale	Administration responsable de la gestion		L'équipe d'évaluation
4ème (a)	Coproduite	Autorités publiques (élus ou Directeurs)	Gestionnaires + Opérateurs	L'équipe d'évaluation
4ème (b)	Pluraliste	Autorités publiques (élus ou Directeurs)	Autorités publiques + Société civile + (Opérateurs)	
5ème	Emancipation	Société civile		

(Source : Baron, Monnier, 2003)

1.5. Sortie des controverses et évaluation technologique et scientifique

En 1972 le Congrès des Etats-Unis inventa une nouvelle forme d'évaluation en se dotant d'un *Office of Technological Assessment* (OTA). Cette organisation poursuivait ainsi sa logique de contrôle de l'exécutif en complétant son dispositif d'évaluation économique et budgétaire par un autre dispositif centré cette fois sur les enjeux de nature scientifique et technique (Moatti, 1983).

L'OTA est un organe du Congrès américain instauré en 1972 par la volonté chez une majorité de parlementaires de disposer, en matière de R&D, d'un potentiel d'évaluation des décisions qui soit indépendant des décideurs industriels et gouvernementaux. Le constat posé par ces parlementaires est celui d'une impuissance ou d'une absence de maîtrise décisionnelle face à la complexité et à l'incertitude des programmes technologiques sur lesquels ils sont amenés à se prononcer » (Lobet-Maris, Kusters, 1992, p.439).

Les principales motivations à la base de la création de l'OTA étaient les suivantes :

- Tout d'abord, la recherche scientifique et technique est depuis un siècle considérée comme un moteur essentiel du développement économique et social, mais aussi comme une composante fondamentale des rapports de puissance (comme en témoigne la compétition entre les USA et l'URSS dans le domaine de l'espace). De plus dès les années 60 les innovations en matière informatique ou en biologie ont montré les bouleversements que pouvaient susciter les résultats de la recherche et ses applications. La question est alors posée d'une possible régulation politique de tels enjeux. Et plus explicitement : est-il possible de laisser en dehors du débat démocratique de telles questions ? Les élus se préoccupaient de l'accaparement d'enjeux majeurs par des structures techno-scientifiques. Le secteur privé et les segments du secteur public engagés dans la recherche peuvent-ils s'autoréguler et ne rendre de compte que très partiels via leurs demandes budgétaires ?
- Ensuite, l'évaluation technologique se veut une réponse à un ensemble de controverses amorcées dans les années 50, mais qui ont pris de la vigueur avec le développement des mouvements de STS (*Science Technology and Society*) au cours des années 60 (Barthe, 2002). Il s'agit d'une mobilisation critique contre la prétention de la science positiviste à une neutralité axiologique et à sa capacité d'accomplir son développement en toute autonomie, selon des rationalités strictement internes. Le mouvement STS, au contraire, a cherché à démontrer comment les valeurs et intérêts sociaux et politiques interfèrent avec la production de connaissances. La question de l'ambivalence des

sciences et des techniques est aussi au cœur de ces raisonnements, c'est-à-dire la mise en évidence aux côtés des aspects positifs de nouvelles incertitudes et parfois de nouveaux risques déjà identifiés (Carson, 1962 ; Ellul, 1964). La question du nucléaire a été fondatrice de ces débats, mais ils ont aussi porté rapidement sur l'impact des produits chimiques (interdiction du DDT en 1969) ou la mise au point de l'avion supersonique civil SST .

- Enfin, les *Technological Assessment* ont été développées pour renforcer la légitimité sociale et politique des activités de recherche et d'innovation en contribuant à une maîtrise des risques qui leur sont inhérents. Ce type d'évaluation a été porté par ce que Moatti a nommé les « avocats de la rationalité » qui cherchent à « concrétiser l'idéal de l'ingénieur multidisciplinaire en mesure de trouver des solutions technologiques aux problèmes sociaux » (1983, p. 216). Dans cette perspective ce sont l'appréciation des risques et l'incitation à la vigilance et à la sécurisation des systèmes qui concentrent l'essentiel de l'attention : « La démarche se fera aussi très normative et quantitative et visera à définir des normes de sécurité, des seuils de risque acceptables par l'introduction du calcul économique dans la pondération des coûts liés à l'accident et de ceux relatifs aux mesures de sécurité à mettre en œuvre. Le débat social sur les développements technologiques disparaît progressivement au profit d'un autre, technico-économique, sur la sécurité des développements ; un débat où se mêlent aussi de savants arbitrages politiques, tant les intérêts en jeu sont importants » (Lobet-Maris, Kusters, 1992, p.438).

Malgré son ancrage dans un rationalisme très classique, l'OTA sera supprimé soudainement en 1996 au bout de quatorze ans à l'initiative des sénateurs républicains (sous le mandat Clinton). Sans surprise, ils invoquent la nécessité de faire des économies budgétaires, mais aussi les réticences fréquentes de l'OTA à l'égard des initiatives des gouvernements républicains antérieurs (Nixon, puis G.H. Bush). Une analyse plus fine a également mis l'accent sur deux autres critiques. Tout d'abord, l'OTA aurait progressivement outrepassé son champ de compétence en prenant des positions de plus en plus morales et éthiques (comme ce fut le cas par exemple pour les biotechnologies). Ensuite, l'OTA n'aurait pas suffisamment joué son rôle de conseiller gouvernemental en développant des points de critiques dépourvus de propositions d'actions concrètes

susceptibles d'aider les décideurs à faire évoluer leurs choix (Mironesco, 1997, p.97 ; Barthe, 2002 p. 36-37).

En Europe, il faut attendre les années 80 pour que le modèle du *Technological Assessment* soit repris. Les premières créations eurent lieu en France (1983) et en Flandre (1984). Mais ce transfert s'accompagne de reconfigurations importantes (Moatti, 1982 ; Petrella, 1992 ; Barthe, 2002). D'un côté, les méthodologies sont diversifiées et l'évaluation se voudra de plus en plus pluri-disciplinaire, accordant une place croissante aux sciences sociales (et pas seulement à l'économie, mais aussi à la sociologie). De plus, sous l'influence du modèle scandinave des Conférences de consensus, la consultation, voire la participation du public ont été progressivement intégrées dans les démarches. Ainsi, en Allemagne, en Belgique et en Espagne, l'évaluation des choix technologiques est une composante importante du dialogue social (Boy, 2000)²⁸. D'un autre côté, les rattachements institutionnels sont variés. Le lien avec les parlements reste cependant une forme fréquente mais non exclusive. Les activités de *Technological Assessment* sont aussi souvent situées dans des instituts ou agences gouvernementales. L'encart n°7 est indicatif de cette diversité.

Le cas de la France est intéressant dans la mesure où s'il est un initiateur en Europe avec la loi du 8 juillet 1983 créant un office parlementaire commun aux deux assemblées (Claret de Fleurieu, 1995), cette décision fait suite à une série de tentatives inabouties entreprises à partir de 1975 (Barthe, 2000, p. 56-58). Y. Barthe considère que trois obstacles ont longtemps entravés les initiatives. Tout d'abord, la puissance toute particulière des « grands corps » d'ingénieurs qui estimaient détenir, sinon le monopole, du moins toutes les compétences requises pour effectuer l'essentiel du *Technological Assessment* tant pour le Parlement que pour le gouvernement. Ensuite, la crainte récurrente des gouvernements successifs de voir ce type d'activité fournir une matière à controverse et polémique politiciennes au bénéfice de l'opposition. Les questions énergétiques et environnementales étant alors perçues comme particulièrement sensibles. Enfin, les réticences d'une partie non négligeable des élus qui redoutaient l'entrée en force dans leur délibération de points de vue d'experts qui auraient contraint leurs choix et

²⁸ En France la conférence de consensus organisée en 1998 par l'OPECST sur les OGM demeure l'exception qui confirme la règle d'une évaluation des choix technologiques effectués par des experts en milieu relativement confiné (Boy, 2000).

substitué des critères rationalistes aux normes et valeurs démocratiques. Retenons au passage que ces trois obstacles à l'évaluation ne sont pas propres aux domaines scientifiques et techniques. Ils ont en France une portée beaucoup plus générale. L'alternance de 1981 et la sensibilité d'une partie des élus socialistes aux enjeux énergétiques et environnementaux ont créé une fenêtre d'opportunité pour la création de l'OPECST. Cependant, le facteur déclencheur semble avoir surtout résidé dans les initiatives du ministre de la recherche J.P. Chevènement qui organise un grand colloque national (janvier 1982) et veut faire de la science et de la technologie un facteur de développement national. L'Office sera une des composantes de ce projet.

Ces structures d'évaluation ont très tôt été l'objet de débats, non seulement sur leur capacité à orienter véritablement les décisions politiques, mais plus fondamentalement en raison d'un certain degré d'ambiguïté de leur projet qui mêle démocratie et technocratie : « il s'agit d'un côté de donner aux parlementaires une maîtrise sociale sur les politiques technologiques et scientifiques et de l'autre, d'asseoir cette maîtrise sur une expertise qui se veut neutre et scientifique, sortant de la sphère politiques un débat qui lui est intimement lié... l'OTA se situe bien dans le paradigme paradoxal d'une décision dont on voudrait consolider l'assise démocratique par un regain de technocratie » (Lobet-Maris, Kusters, 1992, p.439). Un des initiateurs du *Technological Assessment*, Harvey Brooks²⁹, déclare dès 1973 :

« Le mouvement en faveur de la création de mécanismes d'évaluation de la technologie est sans doute fondé sur l'idée erronée que seuls des experts sont capables d'opérer les choix difficiles qui s'imposent. Dans le système politique américain, en particulier, on est très tenté de recourir à l'analyse « objective ou scientifique » pour légitimer le consensus politique et écarter les critiques qui peuvent être déguisées en avis d'experts, mais qui représentent généralement un ensemble de postulats axiologiques différent de l'opinion prédominante » (1973, p.276)».

²⁹ Brooks H., physicien et ingénieur, responsable du département de physique appliquée de Harvard. Après avoir fait partie du President's Science Advisory créé en 1957, déçu des relations avec l'équipe présidentielle se tourne avec un groupe de collègues vers le Congrès et ils seront les inspirateurs de l'OTA. Il est l'auteur d'un célèbre rapport pour l'OCDE en 1971 « Science, croissance et société ».

Encart n° 7 : L'institutionnalisation du *Technological Assessment* en Europe

Institutions créées à l'initiative des parlements		
France	1983	OPECST
Danemark	1986	DBT
Pays Bas	1986	Ratheneau Institut
Union Européenne	1987	STOA
Allemagne	1989	TAB
Grande Bretagne	1989	POST
Commissions parlementaires permanentes		
Finlande	1986	FCF
Grèce	1997	GPC – TA
Italie	1997	VAST
Missions à long terme dans des organismes publics		
Allemagne	1987	ITAS
Suisse	1994	Bureau TA – Publiforum
Autriche	1995	ITA
Union Européenne	1997	ESTO
Institutions régionales		
Flandres	1984	STV
Baden Wütemberg	1991	AT
Rhénanie Westphalie	1992	AKTAB
Wallonie	1996	Commission TA, Conseil wallon de politique scientifique

(Source : La Lettre EMERIT, 24, mai 2000)

Une autre ambiguïté est relevée par Barthe (2000, p.64) à propos de la création de l'OPECST français : « Cette unanimité (en faveur de la création de l'Office), au sein du Parti socialiste notamment, ne se réalise ... qu'au prix d'une certaine ambiguïté concernant le rôle qu'est censé jouer cette instance : s'agit-il à travers cet organisme d'évaluer, de contrôler, éventuellement de critiquer les décisions en matière scientifique et technique prises par le gouvernement, ou s'agit-il d'encourager et de stimuler le développement technologique ? ». Ainsi le ministre souhaitait que l'OPECST « soutienne » les innovations scientifiques et techniques, mais ne voulait pas entendre parler de « maîtrise » car le terme sous-entendait une distance, voire un scepticisme à l'égard de l'optimisme scientifique. Comme dans le cas de l'Office américain, nous voyons à propos de l'OPECST que la crainte d'une utilisation des résultats de ce type d'évaluation au profit d'une critique des choix scientifiques et technologiques est une constante. Le *Technological Assessment* a été conçu et demeure comme un moyen important de légitimation des activités de recherche. Il devait faire obstacle aux positions critiques dites « anti-science ». Sa fonction principale a donc toujours été beaucoup plus un moyen permettant une meilleure appropriation des enjeux scientifiques et techniques par les décideurs politiques qu'une évaluation de l'opportunité et des risques des choix effectués en ce domaine

1.6. Le retour de la performance publique

A partir des années 2000, après l'échec du CNE, la question de l'évaluation n'est plus posée en termes de modernisation de l'action publique, ni d'organisation de contre-pouvoir, mais est couplée à la dynamique de recherche d'une plus grande performance publique. Cet objectif n'est pas nouveau puisqu'il s'inscrit dans un courant beaucoup plus long de réforme de l'Etat (Bezes, 2009). Mais le rattachement récurrent, dans cette période, de l'évaluation dans le pilotage des différents chantiers de modernisation de la gestion publique constitue un fait inédit.

L'évaluation est intégrée à un ensemble d'instruments et de réseaux d'expertise qui se disputent l'influence dans le domaine de la réforme de l'État (Bezes, 2009). Elle est dès lors mobilisée de différentes manières dans les deux grands chantiers successifs de réforme entrepris à partir de 2002, à

savoir la mise en place de la Loi organique relative aux lois de finance (LOLF) et le lancement de la Révision générale des politiques publiques (RGPP). On verra cependant qu'au-delà de la revendication d'une approche soit disant « évaluative », les décideurs ont favorisé des démarches qu'on peut, d'un point de vue technique, considérer comme voisines de l'évaluation (audits, contrôle de gestion, inspection...) mais qui l'ont, en quelque sorte, « remplacée » ou assignée à des tâches spécifiques (Barbier, 2010).

L'évaluation des performances et des budgets : la LOLF

L'adoption, en août 2001, puis la mise en œuvre systématique de la LOLF a contribué à refondre l'ensemble de la procédure budgétaire. Cette réforme 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.

A l'origine, les premiers initiateurs laissaient entendre que la LOLF allait satisfaire le besoin d'évaluation de politique publique, ce qui n'a pas été le cas, comme le prouve en particulier la persistance structurelle du financement d'évaluations sectorielles dans de nombreux départements ministériels. Pourtant, considérée avec le recul de la longue durée, la dynamique générale ainsi enclenchée est beaucoup plus puissante que celle qui s'attache à la pratique spécifique de l'évaluation (Bezes, 2009).

A certains égards, les espoirs fondés à l'origine dans la LOLF n'étaient pas usurpés. Dans son principe, cette dernière oblige les administrations à déterminer des objectifs et des programmes par rapport auxquels sont attribués les crédits budgétaires. Si ce texte a pour visée première un renforcement de la portée de l'examen budgétaire par les Assemblées, il accorde une place explicite à l'évaluation des performances en lui donnant une valeur constitutionnelle (la LOLF est une loi organique), en élargissant son champ à toutes les actions relevant des finances publiques et en en faisant un complément des activités classiques de contrôle. La LOLF demande également aux acteurs publics de déterminer des objectifs d'action, de définir les indicateurs qui permettent d'en suivre la réalisation et de mettre en place des systèmes d'information finalisés. Les deux lacunes

chroniques des politiques publiques françaises devaient ainsi être, sinon résolues, du moins en partie comblées : la détermination des objectifs et une information précise sur les actions et les résultats. Ainsi l'article 51 prévoit que le projet de loi de finances comporte des annexes « accompagnées du projet annuel de performances de chaque programme précisant : la présentation des actions, des coûts associés, des objectifs poursuivis, des résultats obtenus et attendus pour l'année à venir mesurés au moyen d'indicateurs précis dont le choix est justifié ». L'objet n'est pas ici la discussion d'ensemble de ce nouvel instrument de suivi de l'action publique, mais de s'interroger sur sa capacité à intensifier les activités d'évaluation. Deux limites peuvent être pointées pour juger de la capacité de la LOLF à parer aux insuffisances des pratiques évaluatives. Il s'agit tout d'abord d'un instrument de comptabilité publique et sa pertinence repose en grande partie sur la validité des nomenclatures qui ont été élaborées. Comme l'indiquent les parlementaires qui ont anticipé une partie des biais possibles : la logique de la LOLF « pourrait s'évaporer si la définition pratique des missions et des programmes devait déboucher sur une nomenclature budgétaire ne se conformant pas à la volonté exprimée par le législateur organique. Tel serait en particulier le cas si une logique institutionnelle devait prévaloir sur une approche fonctionnelle, assimilant mission à ministère et programme à direction administrative »³⁰. On peut élargir l'interrogation à l'ensemble des catégorisations utilisées, quels agrégats recouvrent-elles, quelle homogénéité présentent-elles entre elles et par rapport à celles élaborées par d'autres secteurs administratifs ? Un élément de réponse peut être recherché dans l'analyse des auteurs de ces nomenclatures et celle des procédures qu'ils ont utilisées pour les définir et les valider. On aurait pu penser que l'élaboration des grilles d'activités et des indicateurs de performance offrait une excellente opportunité pour une auto-évaluation des objectifs, moyens et résultats des secteurs administratifs. Un temps d'objection des pratiques en cours, d'analyse des activités et de leurs produits aurait pu susciter, dans le même temps, un apprentissage collectif et la négociation de nouveaux objectifs et moyens de suivi. Pour reprendre une terminologie classique, une telle évaluation en amont de la réforme aurait pu être à la fois endo-formative (regard critique sur ce qui est à l'œuvre) et normative (détermination de buts

³⁰ Rapport du Sénat n°388 du 9 juillet 2003 : *Mise en œuvre de la LOLF : un outil au service de la réforme de l'Etat*, Jean Arthuis, Commission des finances.

renouvelés). Pour cela, il aurait fallu que la mise en œuvre de la LOLF se fasse « par le bas » (*bottom up*), en s'appuyant d'abord sur la compréhension que les acteurs opérationnels de l'action publique ont de leurs actions et des nouvelles orientations qu'ils peuvent lui donner. Cela aurait exigé autant de volonté politique que de temps pour donner du sens à l'exercice. Il n'en a pas été ainsi, et même si le projet de LOLF était en gestation depuis plusieurs années³¹.

Deux réserves principales sont aujourd'hui formulées sur ce nouvel instrument : l'un relatif à ses conditions d'élaboration, l'autre relatif à sa technique propre. D'un côté, l'entrée en vigueur s'est faite à marche forcée et selon une dynamique hiérarchique descendante (*top down*) accompagnée d'une concertation a minima. Dans la quasi-totalité des ministères se sont les services financiers qui ont été en charge de la réalisation des outils de la LOLF et ils ont accompli leur ouvrage de façon assez isolée en utilisant les apports de certains services de façon plutôt instrumentale (fourniture d'indicateurs pour des catégories d'action prédéfinies..). D'un autre côté, la logique d'agrégation pour l'élaboration des objectifs s'est faite davantage à partir de critères organisationnels qu'en fonction de finalités concrètes. Quant aux critères de choix des indicateurs de performance, ils nous informent davantage sur ce qui est aujourd'hui mesurable (on est parti des données déjà existantes) que sur les variables les plus pertinentes pour rendre compte des objectifs affichés, mais qui seraient en général à produire (Lacaze, 2005).

Au bout du compte, cette rationalisation de la présentation budgétaire ne semble pas susceptible de se substituer à l'évaluation. Elle devrait clarifier les affectations financières, permettre une certaine appréciation de la performance, mais l'instrument ne paraît pas apte à aider à la compréhension des dynamiques d'action et des raisons pratiques qui les gouvernent ...

³¹ Le projet trouve son origine en 1998 avec l'initiative de parlementaires, notamment le député socialiste Didier Migaud alors rapporteur général de la commission des finances, dans la création d'un groupe de travail sur l'efficacité de la dépense publique et le contrôle parlementaire. Laurent Fabius, alors président de l'Assemblée nationale, défend l'idée d'un renforcement des pouvoirs de contrôle du Parlement. L'ensemble débouche sur un rapport fondateur (Rapport du groupe de travail sur l'efficacité de la dépense publique et le contrôle parlementaire. Contrôler réellement pour dépenser mieux et prélever moins. Assemblée nationale, 27 janvier 1999). Le retour de L. Fabius au ministère de l'Economie et des Finances, en mars 2000, favorise l'inscription de la réforme à l'agenda autour d'une stratégie d'alliances entre le gouvernement, l'Assemblée nationale, le Sénat et la direction du Budget.

Les audits de modernisation et la RGPP

Parallèlement à la mise en place de la LOLF, une autre dynamique se développe sous l'impulsion des audits de modernisation et du lancement de la Révision Générale des Politiques Publiques.

En juin 2005, dans le nouveau gouvernement Villepin, le portefeuille de la « réforme de l'État » est significativement attribué au ministre en charge du budget, Jean-François Copé. Ce rattachement constitue une rupture historique avec l'attribution traditionnelle de la réforme administrative au ministère de la fonction publique qu'avait, en quelque sorte, renforcé et institutionnalisé les mesures liées au « Renouveau du service public » en 1989. En juillet 2005, la création de la direction générale de la modernisation de l'État (DGME), qui dépend du ministère en charge du budget, achève de déposséder les autres organismes auparavant en charge de la réforme de l'État.

Dans ce cadre, le ministre du budget et de la réforme de l'État, opposé aux méthodes aveugles « de la hache ou du rabot », lance en septembre 2005 une « phase industrielle d'audits de modernisation » dans le but d'« utiliser plus efficacement l'argent public ». Les données officielles recueillies dans le cadre d'un rapport d'inspection paru en septembre 2012 indiquent qu'au terme des sept vagues d'audits menées entre 2005 et 2007, 170 audits ont été réalisés sur les sites des ministères couvrant près de 150 milliards d'euros de dépenses³².

La stratégie de réforme de l'État à cette époque prend aussi la forme d'un vaste dispositif dit de « révision générale des politiques publiques » : la RGPP. Dans ses objectifs, la démarche revendique explicitement une inspiration tirée d'expériences étrangères antérieures, menées au Canada (*Canadian Program Review* dans les années 1995-1996) et en Grande-Bretagne (*Comprehensive Spending Reviews* depuis 2002). Le principe affiché de la RGPP est le suivant : faire examiner les objectifs, les dépenses, les résultats et les *modus operandi* des grandes politiques publiques ministérielles par des équipes d'audit. Son pilotage regroupe les services de la DGME ainsi que des membres du Secrétariat d'État à la Prospective, créé

³² IGA, IGF, IGAS (2012), « Bilan de la RGPP et conditions de réussite d'une nouvelle politique de réforme de l'État », Rapport remis le 12/09/2012.

en 2007 et dont l'attribution comporte explicitement « l'évaluation des politiques publiques » (ce Secrétariat d'Etat sera réformé en 2009 et la mission d'évaluation des politiques publiques confiée au Ministère du Budget). Le questionnement est large et susceptible de remettre en cause les institutions les plus robustes. Une focalisation sur les enjeux budgétaires et le souci de redimensionner l'État en recourant, chaque fois que nécessaire, à l'externalisation et au secteur privé sont exprimés. L'approche qui dépasse les questions d'optimisation des structures et de promotion d'une administration plus innovante possède certaines caractéristiques de l'évaluation de politiques publiques dans son questionnement. La grille d'analyse formalisée en septembre 2011 repose ainsi sur sept questions clés, à l'époque bien perçue par les fonctionnaires, qui aspirent de longue date à une revue de leurs missions et de leur adéquation avec les moyens disponibles. D'un point de vue évaluatif ces questions interrogent à la fois la pertinence (que faisons-nous ? quels sont les besoins et les attentes collectives ?), la cohérence (qui doit le faire ? Qui doit payer ?), l'efficacité (faut-il continuer à faire de la sorte ?) et l'efficience (comment faire mieux et moins cher) des choix et des actions. Elles sont à la fois rétrospective et prospective (quel scénario de transformation).

Si la RGPP, à l'origine, a largement puisé dans le vocabulaire de l'évaluation, après cinq années de mise en application, il est clair désormais qu'elle n'a pas plus que la LOLF effectivement pénétré sur les terres de l'évaluation des politiques publiques, dans le sens donné par le milieu international des évaluateurs

Le rapport des trois inspections mentionné précédemment montre ainsi que la RGPP a souffert du point de vue de ses méthodes :

- le rythme de la RGPP et son horizon temporel, marqué par la volonté d'obtenir très rapidement des résultats se sont avérés difficilement compatibles non seulement avec l'exercice de concertation mais aussi avec la préparation méthodologique requise par l'évaluation de la performance. Le rapport cite par exemple le cas des audits réalisés sur l'Agence pour la cohésion sociale et l'égalité des chances (Acsé) dont les méthodes d'évaluation proposées par le guide RGPP conçu par l'IGF, avec l'appui de l'IGAS et du CEGEFI, n'ont pas permis d'apprécier réellement l'efficacité des actions menées.

- du point de vue des conditions de mise en œuvre, certaines mesures (l'adoption des mesures de simplification, l'unification des guichets et des procédures ou le développement des télétraitements) n'ont pas toujours été accompagnées d'une évaluation en temps utile des coûts induits, des effectifs et compétences nécessaires ni des investissements informatiques requis.
- le faible degré de concertation et de mise en débat des choix et des opérations réalisés constitue également une limite de l'exercice. Comme l'indique un autre rapport parlementaire remis fin 2011 par F. Cornut-Gentille et C. Eckert³³, la RGPP a « souffert de n'être relayée que par des rapports publics anonymes incapables d'engendrer un réel débat public ». Le Parlement a été absent au moment de la prise de décision et son information limitée dans la phase de mise en œuvre.
- le rapport des inspections conclut enfin à la nécessité d'inscrire dans la durée et dans un processus continue la rénovation de l'action publique en sortant des « démarches évènements successives » (p.63). Cela suppose que la nécessité des évaluations à terme soit inscrite dès le départ dans le processus de réforme (bilan *ex ante*, affichage des objectifs, identification d'indicateurs...). Le rapport mentionne à cet égard que plus secrétaires généraux rencontrés dans le cadre de ce bilan se disent favorables au renforcement des fonctions d'évaluation au sein des administrations.

La mise en œuvre de la LOLF et de la RGPP attestent donc que ces deux chantiers de modernisation ont eu le souci de mobiliser les finalités et les outils de l'évaluation, en lien avec la recherche d'une plus grande efficacité de l'action publique. Cependant, l'un comme l'autre de ces dispositifs a davantage utilisé le vocabulaire et les enjeux de l'évaluation qu'il n'en a réellement appliqué la méthodologie et les principes opérationnels. En raison des délais mais aussi des buts poursuivis (la recherche rapide d'économies budgétaires), les techniques et procédures mobilisés se sont apparentés à des démarches voisines (pour ne pas dire étrangères) à l'évaluation de politiques publiques : audits, inspections... En écartant le Parlement comme les

³³ Assemblée Nationale, Comité d'évaluation et de contrôle, « Une évaluation de la RGPP : méthode, contenus, impacts financiers », *Rapport d'information*, n° 4019, enregistré le 1^{er} décembre 2011.

opérateurs des politiques publiques (au profit de consultants du secteur privé), le pilotage de ces deux dispositifs n'a permis non plus de créer les conditions favorables à la mise en place d'une modernisation de l'administration basée sur l'évaluation de ses activités.

PARTIE 2 : L'organisation, les pouvoirs, les finalités

« En définitive, je préfère parler de la compétence des évaluateurs au quotidien – sujet perfectible et factuellement cernable – plutôt que de traiter de la place de l'évaluation dans la démocratie au début du 21^e siècle – utopie sympathique et qui n'engage que qui veut bien l'écouter » (J.Cl. Thoenig, Société française d'évaluation, Rennes, 15 juin 2000).

Depuis les années 90, trois tendances générales s'observent dans les démocraties occidentales : une évolution des Cours des comptes du contrôle de légalité vers l'évaluation, le développement d'activités d'évaluation au sein des parlements, enfin et plus récemment, la mise en place d'activités d'évaluation au niveau territorial.

2.1. L'évolution de toutes les inspections et contrôle des comptes vers l'évaluation

Nous avons présenté au début de la première partie le contexte général de renforcement d'une rationalité économique appliquée à l'action publique.

Cette dynamique va avoir un effet important sur l'évolution des « Cours des comptes » ou organisme équivalent. C'est tout d'abord en Grande Bretagne que cette nouvelle orientation est concrétisée en 1960 avec la réforme du *National Audit Office* (créé en 1860). Une activité évaluative est introduite sous le nom de « *Value for money studies* » qui repose sur la définition et l'application d'indicateurs de performance censés mesurer les impacts de l'action publique (encart n°8). Ceux-ci sont alors saisis principalement sous l'angle de leur effectivité (ce qui a été mis en place concrètement ou pas), celui de leur efficacité (quels effets de changement et quels résultats ont été obtenus), enfin, celui de leur efficience (une action publique à quels coûts ? Mise en relation des financements engagés avec les résultats, les gains quantitatifs et qualitatifs). Sur la base de ces travaux, le *National Audit Office* effectue également un travail de conseil pour le gouvernement et le Parlement en vue de la conduite de réforme. Une réforme intervenue en 1983

a élargi ce type de compétences à l'évaluation des politiques locales dans un contexte où la « culture de l'audit » a été inculqué à toutes les activités sociales (Faucher-King, Le Galès, 2010, p. 89-94). Aujourd'hui le NAO emploie de l'ordre de 1 000 personnes et effectue annuellement 600 audits financiers et 55 « *Value for money studies* » (rapports approfondis par secteurs) (encart n°8).

Un séminaire franco-britannique sur l'évaluation de politiques publiques, organisé par le Conseil scientifique de l'évaluation et le *Tavistock Institute* s'est tenu à Paris en janvier 1998³⁴. Par comparaison, la situation de l'évaluation en Grande Bretagne apparaît comme fortement institutionnalisée. Ces activités y sont plus développées, politiquement soutenues, réparties à tous les niveaux de l'administration et dotées en budgets et en personnels. L'approche demeure focalisée sur le contrôle de la performance, les analyses compréhensives de la mise en œuvre des programmes est très secondaire. Une des originalités du système anglais est la possibilité pour le National Audit Office de répondre à des demandes émanant de clients privés, en particulier les Fondations œuvrant dans le domaine de la santé. Enfin, certaines « *Value for money studies* » sont menées en collaboration avec des bureaux d'expertise externes qui contribuent surtout sous l'angle organisationnel et plus qualitatif.

Aux Etats-Unis, c'est en 1974 qu'intervient une réforme du *General Accounting Office* (créé en 1921) qui ajoute l'évaluation aux compétences classiques de vérification de la régularité des comptes publics. Comme cela a été indiqué plus haut à propos des méthodes PPBS³⁵, les démarches évaluatives ont été développées aux Etats-Unis, plus qu'ailleurs comme une aide à la décision parlementaire et un contre-pouvoir vis-à-vis de l'exécutif. Une seconde réforme est intervenue en 1998 qui renforce les pouvoirs et compétences de cet organisme qui compte aujourd'hui de l'ordre de 5 000 personnes réalisant près de 4 000 rapports par ans. Ainsi, les parlementaires disposent en permanence d'une information extrêmement poussée sur l'ensemble des activités fédérales.

³⁴ CEPP, *Séminaire franco-britannique sur l'évaluation de politiques publiques*, CSE - Tavistock Institute, 29-30 janvier 1998.

³⁵ Cf. *supra* p. 16 in 1.1

Encart n° 8 : What is a value for money study ?

A value for money study focuses on a specific area of government expenditure, and seeks to reach a judgement on whether value for money has been achieved. We define good value for money as the optimal use of resources to achieve the intended outcomes. Our role is not to question government policy objectives, but to provide independent and rigorous analysis to Parliament on the way in which public money has been spent to achieve those policy objectives.

As well as reaching an overall conclusion on value for money, we make recommendations on how to achieve better value for money and to improve the services under examination. Our reports form the basis of hearings of the Committee of Public Accounts and audited bodies respond to the recommendations made by the Committee on the basis of our reports. In some circumstances, we conduct a follow-up study to measure progress against the recommendations we have made

We aim to be at the leading edge in assessing public service performance, and to adopt innovative approaches and techniques. Typically, a study will use a mix of quantitative and qualitative methods. The methods we commonly use include:

- Financial analysis*
- Analysis of management information*
- Documentary review*
- Interviews or focus groups with departmental and other staff*
- Literature review*
- Surveys of practitioners or service users*
- Benchmarking with other organisations or other countries*

We work to standards that draw on best practice in audit, research and financial management. We use staff with a wide range of professional expertise and bring in specialists from outside the organisation when required. Our studies generally take between three and 12 months from inception to publication.

Our value for money reports are presented to Parliament. Most are considered by the Committee of Public Accounts (PAC) in hearings at which members take evidence from the senior officials of organisations under scrutiny. The PAC then publishes its own report and recommendations, to which the Government must respond.

(Source :

http://www.nao.org.uk/what_we_do/value_for_money_audit/what_is_vfm_audit.aspx)

Aujourd'hui, le GAO soumet plus de 1000 rapports par an au congrès. Ils ont un contenu recommandatif à destination des agences de l'exécutif qui doivent répondre devant le Congrès des décisions prises à la suite des recommandations du GAO. Ce sont ainsi plus des trois quarts de celles-ci qui ont été mises en œuvre en moins de quatre ans, ce qui démontre l'influence du GAO, que beaucoup qualifient de « chien de garde du congrès » face aux pouvoirs forts de l'exécutif. Parallèlement au GAO, l'évaluation *ex ante* des programmes est assurée par le *Congressional budget office* (le CBO) depuis sa création en 1974. Il a renforcé le pouvoir budgétaire du Congrès.

Il faut enfin noter l'existence de l'*Office for the management of Budget* (agence du gouvernement fédéral). C'est en 1981, qu'une décision présidentielle (*executive order*) impose aux agences fédérales, sous le contrôle de l'OMB, d'évaluer les coûts et les bénéfices induits par les mesures réglementaires relevant de leur attribution. A partir de ce pôle fédéral, la pratique de l'évaluation s'est progressivement étendue à tous les niveaux de gouvernements (administrations fédérales, Etats, collectivités territoriales) et aux organisations non gouvernementales.

De plus, aux USA l'évaluation des programmes publics est depuis longtemps ouverte au secteur privé. De très grands organismes comme le *Brookings Institute* ou la *Rand Corporation* sont fortement positionnés sur ce marché. Ils sont en compétition permanente pour obtenir les marchés d'Etat. La réputation méthodologique est l'un des meilleurs arguments de vente de leur « produit ».

En France l'évolution de la Cour des comptes intervient progressivement à compter de 1986 avec la création d'un groupe de réflexion original, le « Club Cambon » (encart n°9) qui regroupe des magistrats modernistes de cette institution participant pour la plupart à l'Association Service Public. Leur but est de « contribuer aux rencontres et aux recherches à l'intérieur de la Cour et des chambres régionales des comptes : son objet, dans le respect du démocratique de l'Etat, est de mettre en commun l'expérience de ses membres pour faire progresser l'efficacité publique ». Ce groupe souhaite contribuer à l'évolution de cette institution d'un contrôle de régularité traditionnel vers des activités évaluatives. Il s'inspire explicitement du modèle étasunien du GAO, mais souhaite aussi pouvoir adapter les activités de la Cour des comptes aux évolutions contemporaines, tant en termes de

structures (décentralisation), qu'en termes de conduite des politiques publiques (performance publique et légitimation démocratique). Pour cela, il s'appuie sur les spécificités de leur organisme : son indépendance facteur de crédibilité, sa stabilité dans le temps, son accès privilégié à de multiples sources d'information (et la possibilité si besoin d'en susciter) et la protection de leur confidentialité, son champ de compétence étendu sur le secteur public et para-public, enfin, sa position charnière entre l'exécutif et le Parlement. Le groupe légitime aussi ses positions en se situant dans une tradition de la Cour et en se référant à un de leurs collègues Bernard Gournay qui dès 1983 s'interrogeait pour savoir dans quelle mesure « l'évaluation était une nouvelle perspective pour les contrôles (...) et si l'évaluation pouvait constituer un nouveau principe de gestion ».

Encart n° 9 : Le Club Cambon pour une Cour des Comptes évaluative

L'association Services Publics (ASP), à travers des réunions, des séminaires ou des petits colloques, va œuvrer à l'acculturation à l'évaluation d'une partie de la haute fonction publique. Cette acculturation se fait sans guère de références aux figures méthodologiquement et institutionnellement centrales que représentaient jusque-là la Direction de la Prévision (DP) et le Commissariat Général du Plan dans le domaine de l'évaluation des politiques publiques. Dans la période de la cohabitation, la délégitimation de la DP et du CGP devient totale en la matière. La DP, emmenée par Patrick Vial et Paul Champsaur, ne veut pas être légitime ; le CGP ne peut plus être légitime.

Un indice de cela réside dans la création le 22 mai 1986 du « Club Cambon », sorte de succursale de l'ASP à la Cour des Comptes, visant à faire souffler un vent de modernisme au sein de la population des magistrats et rapporteurs des juridictions financières du pays. Deux thèmes de réflexion prioritaires ont été retenus par le Club Cambon. Tout d'abord, il s'agit de s'adapter à la décentralisation politico-administrative qui est en cours. Le credo des membres du Club Cambon entre ici en congruence avec l'obsession traditionnelle de l'institution Cour des Comptes quant à la menace supposée grande et permanente du démantèlement de l'Etat à sa périphérie. Ainsi, le Club Cambon agit-il comme le producteur et le diffuseur d'un discours de rappel à l'ordre doctrinaire traditionnel à l'adresse essentiellement des Chambres régionales des Comptes qui doivent avant tout former un réseau homogène, un et indivisible, régulé et unifié par la maison-mère. Pour des raisons évidentes l'action du Club Cambon dans ce domaine est interne. Nettement plus extravertis sont les débats que le Club Cambon choisit d'organiser sur le thème de l'évaluation des politiques publiques. Les discours sur le pluralisme institutionnel et méthodologique nécessaire au développement de l'évaluation qui ont marqué le colloque Nioche-Poinsard et le

rapport Deleau ont frappé les esprits de certains. Ces discours, ainsi que la perte de vitesse du Commissariat Général du Plan, incite certains magistrats de la Cour des Comptes membres de l'ASP à vouloir se concentrer sur un projet de modernisation de leur administration d'appartenance dans le sens de l'adjonction d'une compétence officielle en évaluation. La création du Club Cambon est la première étape du projet. L'organisation d'une grande journée d'étude sur le thème de l'évaluation (en octobre 1987) en est la seconde.

Les membres du Club Cambon identifient bien les atouts dont dispose la Cour des Comptes en matière d'évaluation. Ils constatent au premier chef l'indépendance d'enquête de la Cour par rapport à l'Exécutif (droit d'auto-saisine). En second lieu, la Cour dispose d'un accès privilégié à l'information comptable (pas seulement financière) et qualitative (via la pratique d'audit). Troisièmement, la Cour occupe une position charnière entre exécutif et législatif puisque l'une de ses principales fonctions est « d'assister le Parlement et le gouvernement dans le contrôle de l'exécution de la loi de Finances ». Enfin, le Club Cambon apprécie l'activité quotidienne de contrôle de la Cour selon la partition suivante : « un tiers de contrôle de conformité aux règles de la comptabilité publique, deux tiers consacrés à formuler des jugements ou des conseils qualitatifs d'opportunité des dépenses publiques ». Il serait pertinent de valoriser une partie de ces deux derniers tiers en formant les magistrats de la Cour aux méthodes d'évaluation des politiques publiques.

Ceci étant, le Club Cambon n'est pas un repaire de naïfs. Ses membres savent que leur capacité de mobilisation à l'intérieur de la Cour est limitée par des facteurs structurels comme le turn-over très élevé du personnel, l'extrême segmentation des travaux, la surcharge de travail, etc. En outre, si les règles juridiques de comptabilité publique confèrent aux magistrats de cette juridiction une forte légitimité d'investigation, on ne pourrait en dire autant des méthodes d'évaluation si elles venaient à compléter les prérogatives des magistrats. La perspective de rendre plus tangibles les suites données aux rapports de la Cour par la pratique de l'évaluation constitue un argument culturellement trop « fumeux » pour la plupart des magistrats de la Cour. L'exemple du GAO américain est également invoqué par les missionnaires évaluatifs du Club Cambon, mais précisément la réforme « evaluation-oriented » du GAO au début des années 1970 s'est faite par la voie légale et largement de l'extérieur à l'institution. Aussi le Club Cambon organise-t-il la journée d'étude d'octobre 1987 dans le but non seulement de convertir les moins récalcitrants de leurs collègues, mais également pour donner un signal clair à des institutions telles l'Assemblée Nationale, l'Inspection générale des Finances, la Direction du Budget ou certains ministères qui pourraient en retour passer commande d'évaluations à la Cour des Comptes. Diffuser l'idée d'une Cour des Comptes capable d'évaluer des politiques publiques qu'accrédite le Club Cambon par son action pourrait même à terme convaincre le gouvernement d'habiliter cette nouvelle fonction par la loi.

(Source : Spenlehauer, 1998)

Une évolution des pratiques intervient à partir de 1991 avec la publication de « rapports thématiques » en complément du rapport annuel de la Cour. Ce sont de fait des évaluations de programmes publics dépassant largement la dimension « contrôle » et produisant des informations et des avis tant sur le contenu des programmes que sur les modalités et résultats de leur mise en œuvre.

Une autre étape décisive intervient au milieu des années 2000. Lors de la commémoration des 200 ans de la Cour des comptes, en séance solennelle du 5 novembre 2007, le Président de la République annonce une réforme qui a vocation à faire de la Cour « le grand organisme d’audit et d’évaluation des politiques dont notre Etat a besoin ». Cette annonce débouchera sur la réforme constitutionnelle du 23 juillet 2008 qui donne officiellement à la Cour un mandat d’évaluation de programmes publics au bénéfice du gouvernement et du parlement (encart n° 10).

Philippe Seguin est alors le Premier Président de la Cour et il encourage cette évolution. A l’occasion d’un colloque organisé en 2008 à Marseille pour les dix ans de la « Société française d’évaluation », il présente cette nouvelle mission avec beaucoup de modestie, estimant que son institution est « un acteur de l’évaluation en devenir » et « qu’il ne prétend à aucun monopole ». Il entend aussi bénéficier de l’expérience des Cours étrangères et des travaux d’un « groupe de travail international sur l’évaluation des programmes » qui fonctionne depuis 1992. Depuis 2010 ce réseau a commencé à recueillir des informations sur les modalités et résultats des « bonnes pratiques » afin de les faire circuler. Son souhait est d’inscrire l’évaluation dans le champ plus large de l’évolution des cultures professionnelles liées aux activités de contrôle. Plutôt que d’opposer, comme on le fait couramment, contrôle et évaluation, il propose plutôt d’approfondir les continuités entre : audit-financier, contrôle de régularité, audit-performance et évaluation. Pour lui, le modèle du GAO montre les possibilités d’une complémentarité dans les approches.

Encart n° 10 : Une nouvelle mission de conseil

Article 47 – 2 de la Constitution (3/7/2008)

« La Cour des comptes assiste le Parlement dans le contrôle de l'action du Gouvernement. Elle assiste le Parlement et le Gouvernement dans le contrôle de l'exécution des lois de finances et de l'application des lois de financement de la sécurité sociale ainsi que dans l'évaluation des politiques publiques.

Par ses rapports publics, elle contribue à l'information des citoyens.

Les comptes des administrations publiques sont réguliers et sincères. Ils donnent une image fidèle du résultat de leur gestion, de leur patrimoine et de leur situation financière ».

La cour des Comptes définit ainsi cette nouvelle mission de conseil : « L'évaluation part d'un a priori neutre et cherche autant à valoriser les aspects positifs qu'à critiquer les aspects négatifs d'une politique publique, qu'elle peut aider à conforter ou remettre en cause.

Depuis la loi du 3 février 2011, le président de l'Assemblée nationale et le président du Sénat peuvent saisir la Cour de demandes d'évaluation, sous forme de rapports remis au plus tard un an après la demande »

(Source <http://www.ccomptes.fr/Nous-connaitre/Nos-missions/Cour-des-comptes>).

Les premières évaluations de politique publique sur le nouveau modèle ont été réalisées en 2011 et rendues publiques dans le rapport annuel publié en février 2012. Elles sont au nombre de six. Deux l'ont été à la demande du Parlement et portent sur la médecine scolaire et la politique d'hébergement des personnes sans domicile. Quatre l'ont été à l'initiative de la Cour concernant : la politique d'aide aux bio-carburants, la politique en faveur de l'assurance-vie, la politique des relations de l'administration fiscale avec les entreprises et les particuliers, enfin, celle portant sur la sécurité des navires. En introduction, la Cour rappelle « les particularités » de sa démarche évaluative.

- Elle porte sur un objet spécifique cernable, « sur une politique publique et non sur un organisme ou sur des procédures ».
- C'est une analyse par les effets « centrée sur les résultats ... et leur rapprochement entre d'une part des objectifs visés et d'autre part, des

moyens mis en œuvre ». L'approche est pluri-disciplinaire, la politique « est envisagée dans globalité (composantes économique, sociale, culturelle, etc.) et dans sa diversité, ce qui implique, le plus souvent, des approches multi-disciplinaires et le recours à des expertises extérieures ».

- Elle repose sur une intégration de la pluralité des points de vue, « les principales parties prenantes à la politique évaluée sont systématiquement prises en compte, et, dans toute la mesure du possible, associées à l'évaluation, notamment de façon à élargir les informations disponibles et à garantir la pertinence opérationnelle des recommandations ».
- Elle vise un approfondissement des problématique sous-jacentes, « leur analyse est au cœur de l'évaluation – notamment, pour déterminer si les hypothèses sur lesquelles la politique évaluée est fondée sont vérifiées dans la pratique – en ayant le souci constant de pouvoir avant tout identifier les voies pratiques de progrès ». Une des dernières évaluations réalisées porte sur la politique vaccinale (encart n°11).

2.2. Les Parlements

Les relations entre les Parlements et les activités d'évaluation des politiques publiques sont pour le moins paradoxales, et cela dans tous les régimes politiques où la question a pu être abordée. D'un côté, les Parlements par leur légitimité démocratique sont *a priori* l'organisation la mieux placée pour exercer un contrôle de l'activité de l'exécutif. Mais, en pratique, malgré des tentatives diverses, aucun modèle ne s'est imposé qui permette d'atteindre cet objectif.

Malgré les ambitions affichées, les Parlements ne sont jamais parvenus à exercer réellement cette fonction. Cela ressort très clairement de la journée d'étude organisée au Sénat le 7 avril 1994 et qui a dressé le bilan de ce type d'action dans 6 pays (Allemagne, Belgique, Canada, Etats-Unis, France, Suisse (A. Delcamp et al., 1995). Une étude ultérieure confirme les principales conclusions (Maury, 1998).

Encart n° 11 : Rapport sur la politique vaccinale de la France (février 2013)

La Cour des comptes a rendu public, mercredi 20 février 2013, un rapport sur la politique vaccinale de la France commandé par la commission des affaires sociales du Sénat en application de l'article L.O. 132-3-1 du code des juridictions financières.

Principales conclusions : les objectifs vaccinaux établis dans une perspective quinquennale en annexe de la loi de santé publique du 9 août 2004 ont été définis de manière trop uniforme et leur degré de réalisation a été décevant dans le cadre national comme au regard des comparaisons internationales.

L'ensemble des préconisations définies pour chaque âge de la vie par les autorités sanitaires forme le calendrier vaccinal. Celui-ci s'est considérablement étoffé sous le double effet de la disponibilité de nouveaux vaccins et d'un meilleur ciblage des recommandations, qui sont elles-mêmes de plus en plus denses. Ces recommandations sont souvent émises au terme d'un processus qui ne permet pas d'en garantir la clarté ni parfois la cohérence.

Les conditions de prise en charge par la solidarité nationale des vaccins pourraient être améliorées par une diversification des taux de remboursement par l'assurance maladie, des choix plus éclairés sur les stratégies de prévention à mettre en œuvre et des procédures de fixation des prix qui leur soient propres.

La couverture vaccinale ne touche pas identiquement l'ensemble de la population et certaines catégories sont moins couvertes que d'autres: les populations défavorisées et les adolescents qui ont en moyenne moins de contacts avec la médecine libérale et les habitants de certaines zones géographiques traditionnellement plus réticents à l'égard de la vaccination.

Le renforcement de la communication sur les vaccins est indispensable par une présence plus systématique sur Internet, une meilleure définition des cibles prioritaires, un encadrement plus efficace de la communication des producteurs des vaccins.

Le rapport est assorti de 16 recommandations qui visent à : 1 - renforcer l'efficacité des recommandations vaccinales ; 2 - améliorer la détermination des prix et des taux de remboursements des vaccins ; 3 - combler les lacunes de la couverture vaccinale ; 4 - rendre plus efficace la communication sur la vaccination.

L'évaluation dans les « checks and balances » (Etats-Unis, Grande Bretagne)

Théoriquement, tous les gouvernements démocratiques ont cherché à se doter d'un système de « pouvoirs/contre-pouvoirs ». Cependant, malgré la diversité des contextes institutionnel et politique, aucun d'eux n'est parvenu à se doter de dispositifs qui institutionnalisent véritablement cet équilibre des pouvoirs. Les pays anglo-saxons en raison de leur approche plus pragmatique des choix politiques et de leur défiance historique à l'égard d'une autorité centrale trop forte, semblent mieux dotés. Comme nous l'avons montré au point précédent cela repose essentiellement sur l'existence d'organismes de contrôle extérieurs puissants et légitimes comme le *National Audit Office* en Grande Bretagne ou le *General Accounting Office* aux Etats-Unis. Le développement sensiblement plus avancé de l'institutionnalisation des démarches évaluatives dans ces pays peut être, au moins partiellement, attribué à une tradition politique et institutionnelle qui intègre mieux que d'autres la discussion et la critique des décisions publiques. Mais l'appropriation des travaux de ces organismes d'évaluation par les Parlements n'en est pas moins aléatoire. La tradition d'un défaut de cadre approprié à l'expression directe de la société civile amplifie cette lacune.

Les Etats-Unis comme le Royaume-Uni partagent une même attention à l'existence d'un certain équilibre entre « pouvoirs et contre-pouvoirs » (« checks and balances »). Cette préoccupation se traduit notamment par la vivacité du Parlement en tant qu'institution. Aux Etats-Unis, le Parlement dispose de pouvoirs constitutionnels importants, d'une forte autorité et surtout de moyens administratifs et financiers pour exercer un contre-poids puissant à l'égard de l'exécutif. De plus, la très grande fragmentation des centres de pouvoir au Congrès, ainsi que le poids que représente chacun de ses membres constituent une incitation permanente à l'exercice effectif de leurs prérogatives par les membres du Congrès. Le régime des partis joue également dans ce sens. En effet, au-delà d'un bipartisme apparent, la vie politique aux Etats-Unis est marquée par l'existence de clivages mouvants qui pénètrent à l'intérieur de chaque formation politique et atténuent les effets d'une stricte discipline majoritaire sur l'exercice des mandats au Congrès.

Au Royaume-Uni, les moyens propres du Parlement sont moins développés, celui-ci dispose, par conventions, d'accès importants à nombre de nombreuses ressources extérieures. De façon emblématique, il entretient des relations étroites avec le *National Audit Office*. Dans ce type d'organisation politique, il est normal et même attendu que l'exécutif rende compte de ses actions devant les instances élues. Ainsi, l'existence de « secrétaires parlementaires » au sein des ministères témoigne dans plusieurs pays d'une systématisation des relations entre exécutif et Parlement. Tel est le cas en Allemagne mais aussi au Royaume-Uni. Dans ces cas, les secrétaires parlementaires ont un rang politique élevé qui contraste avec le statut beaucoup plus modeste des « assistants » parlementaires des ministres en France. Cette organisation institutionnelle constitue l'assise d'une relation fonctionnelle permanente et dense. De plus, l'organisation du travail parlementaire et des relations entre le Gouvernement et le Parlement sont souvent plus propices qu'en France à un climat favorable à l'évaluation

Deux schémas d'organisation politique favorisent donc l'institutionnalisation des processus d'évaluation des politiques publiques :

1 - ceux dans lesquels la séparation des pouvoirs entre institutions et l'organisation d'une certaine égalité entre eux poussent à la recherche de compromis (Etats-Unis).

2 - ceux dans lesquels la tradition parlementaire est restée vivace en dépit d'un contexte institutionnel ou politique pouvant fonder une forme de suprématie de l'exécutif (Royaume-Uni).

A l'inverse de ces deux pays, en France, les rapports sont restés longtemps très distendus entre le Parlement et la Cour des comptes.

En France : des moyens parlementaires importants pour une évaluation limitée

Les procédés traditionnels de contrôle parlementaire sont nombreux mais ils n'ont jamais abouti au développement d'activités évaluatives systématiques. A partir de sa mission de contrôle, le Parlement a toujours fait de l'évaluation à la façon de M. Jourdain, sans forcément le savoir. Le contrôle parlementaire, parce qu'il implique l'examen d'une situation, d'un fait ou d'un problème donnés et un jugement à son égard, implique une évaluation. Mais l'évaluation au sens plein du terme va au-delà du contrôle.

C'est pourquoi les moyens parlementaires traditionnels apparaissent mal adaptés à cette activité.

- Les questions parlementaires : par le biais de ces questions des demandes sont adressées au gouvernement sur ses actions et leurs résultats. Mais toutes les questions ne donnent pas lieu à une réponse (cas minoritaires) et surtout les délais de réponse (2 mois) ne sont pas toujours respectés. Quant à la valeur des réponses elle est très variable, beaucoup sont formelles et faiblement informatives.
- Les auditions et les demandes d'information : elles émanent d'organes collectifs comme les commissions permanentes ou spéciales. Les demandes d'information souffrent des mêmes limites que les questions : les parlementaires n'ont aucune garantie ni sur le moment, ni sur la qualité de la réponse. Ensuite rares sont les auditions qui ont pour objet l'évaluation d'une politique à proprement parler.
- Les commissions d'enquête : elles constituent un des leviers parlementaires importants pour contrôler l'action des pouvoirs publics. Mais leur objet est juridiquement contraint et l'évaluation n'entre pas dans ce cadre. De plus, le caractère politique de la composition de ces commissions les transforme souvent en lieux de confrontation d'intérêts.
- Les missions d'information : elles sont un outil de production intéressant et elles pourraient contribuer à la réalisation d'évaluation par la réalisation de premiers diagnostics. Mais la plupart restent à un niveau descriptif faiblement interprétatif. De plus, leur politisation fréquente atténue leur valeur.
- Les rapports d'évaluation prévus par les lois : certains textes de lois prévoient la publication de rapports dont certains ont un caractère évaluatif. On peut distinguer deux cas de figure :
 - Certaines lois (peu fréquentes) prévoient la publication d'un rapport d'évaluation : ainsi la loi du 25 juillet 1994 relative à la politique familiale suscite un rapport annuel d'évaluation de la politique familiale. Egalement la loi du 30 décembre 1991 prévoit un rapport annuel et une commission d'évaluation en charge de sa réalisation (encart n°12)

Encart n° 12 : Loi de 1991 sur les déchets nucléaires

Art. 4. - Le Gouvernement adresse chaque année au Parlement un rapport faisant état de l'avancement des recherches sur la gestion des déchets radioactifs à haute activité et à vie longue et des travaux qui sont menés simultanément pour :

- la recherche de solutions permettant la séparation et la transmutation des éléments radioactifs à vie longue présents dans ces déchets ;*
- l'étude des possibilités de stockage réversible ou irréversible dans les formations géologiques profondes, notamment grâce à la réalisation de laboratoires souterrains ;*
- l'étude de procédés de conditionnement et d'entreposage de longue durée en surface de ces déchets.*

Ce rapport fait également état des recherches et des réalisations effectuées à l'étranger.

A l'issue d'une période qui ne pourra excéder quinze ans à compter de la promulgation de la présente loi, le Gouvernement adressera au Parlement un rapport global d'évaluation de ces recherches accompagné d'un projet de loi autorisant, le cas échéant, la création d'un centre de stockage des déchets radioactifs à haute activité et à vie longue et fixant le régime des servitudes et des sujétions afférentes à ce centre.

Le Parlement saisit de ces rapports l'Office parlementaire d'évaluation des choix scientifiques et technologiques.

Ces rapports sont rendus publics. Ils sont établis par une commission nationale d'évaluation, composée de:

- six personnalités qualifiées, dont au moins deux experts internationaux, désignées, à parité, par l'Assemblée nationale et par le Sénat, sur proposition de l'Office parlementaire d'évaluation des choix scientifiques et technologiques;*
- deux personnalités qualifiées désignées par le Gouvernement, sur proposition du Conseil supérieur de la sûreté et de l'information nucléaires;*
- quatre experts scientifiques désignés par le Gouvernement, sur proposition de l'Académie des sciences.*

- Enfin, de très rares textes de loi ordonnent une évaluation et déterminent leur mode de réalisation. C'est le cas de l'article 52 de la loi du 1er décembre 1998 relative au RMI. Cet article dispose qu' « il est nécessaire de procéder régulièrement à l'évaluation des actions menées dans le cadre des Programmes Départementaux d'Insertion (PDI) [et que] cette démarche peut être conduite dans un cadre ana-

lytique, afin de mieux mesurer l'efficacité des actions entreprises et de recenser les besoins des publics cibles (...) Elle doit également inclure une dimension prospective dont l'objectif sera l'amélioration du dispositif de suivi local des bénéficiaires et la proposition d'actions innovantes en matière d'insertion (...) Ces bilans, accompagnés des dispositions prises à la suite de l'évaluation, feront l'objet d'une transmission à l'administration centrale, afin d'alimenter la réflexion du ministère sur le sujet ».

Suite aux vœux de Jacques Chirac formulés en mai 1995 sur une modernisation du Parlement, et à l'initiative de Philippe Seguin (alors Président de l'Assemblée nationale) deux lois du 14 juin 1996 ont créé deux offices chargés d'évaluation : l'Office parlementaire d'évaluation des politiques publiques et L'Office parlementaire pour l'évaluation de la législation

Dans les deux cas, la notion d'évaluation est clairement distinguée des notions de « contrôle » (vérification de régularité) et d'expertise (analyse technique). L'évaluation est bien comprise comme intégrant une part de jugement, comportant dans son volet a priori une estimation des moyens requis et dans son volet a posteriori une appréciation des résultats et de l'efficacité.

L'Office parlementaire d'évaluation des politiques publiques, commun aux deux assemblées, est composé de deux délégations, constituée l'une à l'Assemblée nationale et l'autre au Sénat. Cet Office a mal fonctionné. Sa composition bicamérale a conduit à des blocages en cas de cohabitation. D'ailleurs, Laurent Fabius, élu Président de l'Assemblée nationale en 1997, a décidé d'emblée la mise en sommeil des deux offices parlementaires, prenant acte de l'impossibilité de les faire fonctionner avec deux assemblées politiquement opposées (Perret, 2008). Par ailleurs, les procédures de lancement des évaluations se sont révélées lourdes à mettre en œuvre. L'Office n'a pu produire qu'un nombre modeste d'évaluations puisque quatre rapports seulement ont été publiés³⁶.

³⁶ Rapport « La politique maritime et littorale de la France », réalisé sous l'égide du sénateur Philippe Marini (6 mars 1998) ; Rapport « Les aides publiques au cinéma français : le prix d'une réussite » réalisé par le sénateur Jean Cluzel (7 octobre 1998) ; Rapport « Les aides publiques aux entreprises en matière d'emploi : bilan et perspectives » réalisé sous la conduite du député Gérard Bapt (28 avril 1999) [Assemblée nationale : n° 1547, Sénat : n° 329, 1998-1999] ; Rapport «

Eu égard au mauvais fonctionnement de l'Office, c'est sans surprise qu'il a été supprimé, dans un relatif consensus, par l'article 94 de la loi de finances pour 2001 (n° 2000-1352 du 30 décembre 2000).

L'Office parlementaire pour l'évaluation de la législation est quant à lui composé de deux délégations, l'une à l'Assemblée Nationale, l'autre au Sénat. La loi du 14 juin 1996 définit ainsi ses missions: "l'Office est chargé, sans préjudice des compétences des commissions permanentes, de rassembler des informations et de procéder à des études pour évaluer l'adéquation de la législation aux situations qu'elle régit. L'office est également investi d'une mission de simplification de la législation." Alors que des griefs s'expriment fréquemment sur le mal français de l'inflation législative.

Si cet Office parlementaire d'évaluation a été jusqu'à présent maintenu, son bilan apparaît pourtant très pauvre. Seuls trois rapports ont été publiés³⁷.

Dans l'un comme l'autre de ces deux exemples, il a pourtant été rappelé que l'évaluation législative constituait le travail permanent des commissions parlementaires.

Après l'échec de ces deux offices, l'évaluation parlementaire va connaître un second souffle à travers les prémisses de la LOLF. Dès 1999, la suppression des deux offices parlementaires est faite au profit d'une Mission d'évaluation et de contrôle (MEC), relevant de la seule Assemblée nationale. Elle reprend les fonctions de l'Office parlementaire de l'évaluation des politiques publiques dans le champ des finances publiques. Placée sous la responsabilité de la Commission des finances de l'Assemblée nationale, la MEC permet à l'Assemblée de jouer un rôle plus important dans l'évaluation des politiques publiques. Elle résulte de la volonté d'accroître la qualité de l'évaluation parlementaire en matière budgétaire. S'appuyant sur l'expertise de la Cour des Comptes, elle a en particulier pour mission d'inciter les administrations à raisonner en termes d'objectifs-résultats-contrôle. En 2004,

Investissements étrangers en France : l'évaluation du dispositif public de promotion » réalisé sous la conduite du sénateur Serge Vinçon, (5 mai 1999).

³⁷ Le Rapport Gélard déposé le 15 juin 2006 sur les autorités administratives indépendantes ; Le Rapport Hyst déposé le 5 décembre 2001 sur la législation applicable en matière de prévention et de traitement des difficultés des entreprises ; Le Rapport Abertini déposé le 6 mai 1999 sur l'exercice de l'action civile des associations.

le dispositif a été complété par la création d'une Mission d'évaluation et de contrôle des lois de financement de la sécurité sociale (MECSS).

Du côté du Sénat, une instance symétrique est créée avec la mise en place du Comité d'évaluation des politiques publiques du Sénat, institué en juin 2000.

Dans son premier rapport d'activité, la MEC dresse un constat sans concession des carences du travail parlementaire :

« Les projets et propositions de loi venant en discussion au sein de notre assemblée font rarement l'objet, ex ante, d'une évaluation, les notes d'impact devant accompagner les projets de loi restant des plus succinctes, quand elles existent. De même, le Parlement ne dispose, en général, d'aucune véritable simulation des projets de réforme fiscale ou touchant aux prélèvements sociaux qui lui sont présentés. [...] Par ailleurs, les lois adoptées font rarement l'objet, ex post, d'une évaluation en termes d'efficacité. La demande de présentation d'un rapport d'évaluation, destiné à faciliter les adaptations des dispositifs législatifs retenus, voire conditionnant la reconduction des mesures arrêtées, n'est en effet pas systématique. En raison de l'insuffisance d'évaluation ex ante et ex post des textes législatifs, l'activité normative des parlementaires semble, pour l'instant, largement déconnectée de ses implications budgétaires ».

L'Assemblée se trouve ainsi privée de tout examen factuel des effets de la dépense en amont, et de toute possibilité de contrôle en aval du vote du budget. La MEC est conçue pour pallier ces carences.

Encart n° 13 : L'évaluation des politiques publiques à l'Assemblée nationale

L'Assemblée nationale a mis en place depuis plusieurs années, dans le cadre de ses prérogatives de contrôle financier, deux missions permanentes dont l'objectif est de veiller à l'efficience de la dépense publique.

L'une, la mission d'évaluation et de contrôle (MEC) est chargée de contrôler l'utilisation des deniers publics ; l'autre, la mission d'évaluation et de contrôle des lois de financement de la sécurité sociale (MECSS) a pour but de vérifier l'application des lois de financement de la sécurité sociale et de procéder à l'évaluation de toute question relative aux finances de la sécurité sociale.

En outre, créé par la réforme du Règlement du 27 mai 2009, le comité d'évaluation et de contrôle (CEC) permet à l'Assemblée nationale de mettre en œuvre la fonction de contrôle et d'évaluation qui lui est désormais explicitement reconnue par l'article 24 de la Constitution. Ce comité conduit des évaluations de politiques publiques et apporte son expertise sur les études d'impact qui accompagnent les projets de loi déposés par le Gouvernement.

La Mission d'Évaluation et de contrôle (MEC)

Inspirée du National Audit Office du Parlement britannique, la mission d'évaluation et de contrôle (MEC) a été mise en place au sein de la commission des finances en février 1999, à la suite des conclusions du groupe de travail sur le contrôle parlementaire et l'efficacité de la dépense publique qui fut l'initiateur de la loi organique relative aux lois de finances (LOLF). Celui-ci préconisait, au terme de ses travaux, la création d'une structure chargée d'entendre les responsables politiques et administratifs sur la gestion de leurs crédits et de mener des investigations approfondies sur des politiques publiques sectorielles.

Cette mission présente la particularité d'être co-présidée par un député de la majorité et un député de l'opposition, ses 16 membres appartenant à la commission des finances et étant désignés par les groupes politiques, à parité entre majorité et opposition. Le président de la commission et le rapporteur général en sont membres de droit. Les autres commissions permanentes peuvent demander à certains de leurs membres d'y participer.

Le choix des thèmes étudiés par la MEC relève du bureau de la commission des finances, ce qui permet d'assurer la coordination avec l'ensemble des travaux de la commission.

La MEC travaille en collaboration avec la Cour des comptes qui est consultée préalablement au choix des thèmes retenus et dont des membres assistent à ses réunions. Un rapport demandé à la Cour en application de l'article 47-2 de la

Constitution ou du 2° de l'article 58 de la LOLF constitue souvent le point de départ de ses travaux.

Ses rapports sont systématiquement confiés à deux, voire trois, députés, ce qui permet d'associer majorité et opposition ainsi que d'autres commissions permanentes, afin de dégager des conclusions consensuelles.

Ses méthodes de travail (essentiellement des auditions, mais aussi des déplacements sur le terrain et des questionnaires adressés aux acteurs concernés) sont celles de l'ensemble des missions d'information, les auditions étant ouvertes au public et à la presse, sauf exceptions, notamment lorsque sont abordées des questions touchant à la défense nationale.

La LOLF confère en outre à la MEC les pouvoirs étendus reconnus aux rapporteurs spéciaux pour convoquer des témoins et se faire communiquer tous documents, sous la seule réserve des sujets à caractère secret (défense nationale, sécurité de l'État, secret de l'instruction, secret médical).

Les conclusions de la MEC (qui délibère à huis clos) sont soumises à la commission des finances afin qu'elle puisse se prononcer sur la publication du rapport. Ses propositions, tournées vers des mesures concrètes d'amélioration des politiques publiques, font souvent l'objet d'un suivi après un an : rapport ou communication en commission. En outre, à la fin de la XIIIème législature, la MEC a publié un rapport sur les suites données aux propositions qu'elle a présentées au cours de ladite législature.

Enfin, en application de l'article 60 de la LOLF, lorsque les travaux de la MEC donnent lieu à des observations notifiées au Gouvernement, celui-ci est tenu d'y répondre, par écrit, dans un délai de deux mois.

Les rapports d'information de la MEC au cours de la XIII^e législature

- 2007-2008 : Le financement des projets d'équipement naval militaire, l'allocation des moyens des universités, la gestion des ressources humaines au ministère de l'Écologie, de l'énergie, du développement durable et de l'aménagement du territoire, l'immobilier de l'État

- 2008-2009 : Le musée du Louvre, le coût des opérations militaires extérieures, notamment sous mandat international, le financement des services départementaux d'incendie et de secours (SDIS), l'évaluation et perspective des pôles de compétitivité

- 2009-2010 : L'enseignement français à l'étranger, le crédit d'impôt recherche, les recettes exceptionnelles de la Défense en 2009 et 2010,

- 2010-2011 : Les externalisations dans le domaine de la Défense,

- 2011-2012 : La soutenabilité de l'évolution de la masse salariale de la fonction publique, le financement des politiques culturelles de l'État par des ressources

affectées, les financements extra-budgétaires de la recherche et de l'enseignement supérieur

La Mission d'Évaluation et de contrôle des lois de financement de la sécurité sociale (MECSS)

La mission d'évaluation et de contrôle des lois de financement de la sécurité sociale (MECSS) a été mise en place au sein de la commission chargée des affaires sociales en décembre 2004

Les rapports d'information de la MECSS au cours de la XIII^e législature

- 2008 : La prescription, la consommation et la fiscalité des médicaments, les affections de longue durée*
- 2009 : Le bilan de la prestation d'accueil du jeune enfant, la lutte contre la fraude sociale*
- 2010 : Le fonctionnement de l'hôpital*
- 2011 : La lutte contre la fraude sociale*
- 2012 : La prévention sanitaire*

Le comité d'Évaluation et de contrôle (CEC)

Créé par la réforme du Règlement du 27 mai 2009, le comité d'évaluation et de contrôle (CEC) permet à l'Assemblée nationale de mettre en œuvre la fonction de contrôle et d'évaluation qui lui est explicitement reconnue par l'article 24 de la Constitution. Le CEC est une instance de contrôle opérationnel qui, d'une part, conduit des évaluations de politiques publiques, et, d'autre part, apporte son expertise sur les études d'impact qui accompagnent les projets de loi déposés par le Gouvernement.

Travaux du CEC au cours de la XIII^e législature

- 2009-2010 : La mise en oeuvre de l'article 5 de la Charte de l'environnement relatif à l'application du principe de précaution, la politique d'aide aux quartiers défavorisés, l'efficacité des autorités administratives indépendantes*
- 2011 : L'évaluation de la performance des politiques sociales en Europe, la révision générale des politiques publiques, la médecine scolaire,*
- 2012 : Les incidences de la stratégie de Lisbonne sur l'économie française, les territoires ruraux, le service public de l'hébergement et de l'accès au logement des plus démunis, l'aide médicale de l'État, l'évaluation des dispositifs de promotion des heures supplémentaires.*

Les deux structures parlementaires, la MEC et le Comité d'évaluation, ont été créées sans texte, mais ont été consacrées, d'une certaine façon, par

les articles 57 et 59 de la loi organique du 1er août 2001 relative aux lois de finances (LOLF).

- Art. 57 : missions et prérogatives de contrôle des commissions des finances du Parlement.
- Art. 59 : instauration d'une procédure permettant d'obtenir la cessation des entraves exercées à l'encontre des missions de contrôle et d'évaluation du Parlement.

Avec la mise en place de la LOLF, l'évaluation parlementaire semble enfin trouver sa place. La LOLF a réformé en profondeur l'ordonnance du 2 janvier 1959 en assouplissant les règles contraignant les gestionnaires. Les budgets sont dorénavant votés par programmes basés sur des objectifs stratégiques précis.

Le contexte présidant à la création de la MEC et à la LOLF reprend l'objectif d'allocation des crédits budgétaires sur la base d'un examen de l'efficacité comparée de différentes actions. Du point de vue de l'évaluation, l'innovation majeure est le vote de crédits par missions constituées d'un ensemble de programmes, en lieu et place de crédits votés par ministère et par nature de dépense.

Par ailleurs, les gestionnaires, plus autonomes, peuvent alors, au sein d'un programme, réaffecter les crédits (principe de fongibilité) comme bon leur semble avec toutefois l'impossibilité d'augmenter les crédits consacrés au personnel. Des indicateurs de résultats et de qualité viennent évaluer les programmes, permettant de justifier ou d'ajuster les sommes affectées aux différentes missions.

En 2004, un rapport sénatorial rappelait néanmoins que « la LOLF ne permettra pas de faire l'économie d'une véritable institutionnalisation de l'évaluation, d'autant que les « publics » de l'évaluation des politiques publiques sont, par nature, plus diversifiés que ceux de la LOLF ». En 2006, le rapport Brunetière, invitait à son tour à relativiser l'idée que la LOLF pouvait constituer en elle-même un mécanisme d'évaluation. La plupart des indicateurs de résultats associés au programme ne renseignent pas sur la pertinence et la réussite des actions, et pas davantage sur l'efficacité des services.

Si la LOLF ne peut prétendre incarner l'évaluation parlementaire, un rapport sénatorial sur sa mise en œuvre déposé en avril 2006, montrait néanmoins qu'elle avait permis l'émergence d'« un vrai pouvoir d'arbitrage » du Parlement, et par là même a aidé à la revalorisation de l'institution parlementaire et au rééquilibrage des pouvoirs législatif et exécutif. Un renforcement qui tient aussi bien au rôle renforcé du Parlement dans le débat budgétaire qu'en matière de contrôle de la dépense budgétaire et d'évaluation des politiques publiques.

En définitive, les incertitudes quant à la portée évaluative de la LOLF ajoutée aux constats d'échec des expérimentations d'offices parlementaires témoignent d'un bilan assez mitigé de l'évaluation parlementaire en France, encore accentué par la comparaison avec certains de ses homologues étrangers, en particulier le Congrès américain. Seul l'Office parlementaire d'évaluation des choix scientifiques et technologiques semble avoir fourni un travail globalement apprécié des parlementaires. Plutôt que les Offices, c'est la multiplication des dispositions législatives prévoyant qu'un rapport d'évaluation des lois concernées soit transmis au Parlement, et surtout la mise en œuvre de la LOLF qui ont fait vivre la fonction évaluative du Parlement français. Pour autant, comme nous le verrons ultérieurement, cette dernière est loin d'avoir clos les discussions sur la place de l'évaluation au sein du Parlement (cf. conclusion et encart n°20).

2.3. Les activités d'évaluation au niveau territorial

A l'échelle régionale et locale, l'évaluation a été développée selon des mécanismes et des approches différents de ce qui a prévalu au niveau central. Deux courants ont été à l'origine de l'institutionnalisation de l'évaluation au niveau décentralisé. D'un côté, la mise en œuvre des politiques contractuelle, principalement les fonds structurels européens et les contrats de plan Etat-Région sera l'occasion de consacrer l'obligation d'évaluer les programmes contractualisés. De l'autre, les avancées produites par les premières expérimentations locales conduites en dehors de tout cadrage national ont permis la diffusion de pratiques d'évaluation au sein des collectivités territoriales. A travers plusieurs exemples, nous verrons également que l'évaluation a pris part dans un contexte plus large de territorialisation de l'action publique, où elle a été posée comme une source de production de savoirs et de légitimation pour les instances décentralisées et a permis d'orienter une réflexion concourant à une meilleure gouvernance

dans des jeux complexes d'influences réciproques entre les acteurs et à différents niveaux de pouvoirs (Jacob, Varone, Genard, 2007).

L'institutionnalisation sous l'effet des fonds structurels et des contrats de Plan Etat-Région

L'institutionnalisation de l'évaluation dans les politiques régionales est étroitement liée au développement des fonds structurels (FEDER, FSE, FEOGA) à partir du milieu des années 1990. Avant cette date, des pratiques d'évaluation avaient d'abord été développées dans le cadre de la politique de recherche scientifique et technologique communautaire initiée dans les années 60 et structurée autour de programmes pluriannuels. Dans le champ de l'économie régionale, certains pays comme l'Angleterre et les Pays-Bas avaient également mis en place dès les années 70 une tradition d'études empiriques destinées à mesurer les impacts économiques de certaines politiques régionales comme l'aide aux entreprises, mais ces approches ont connu un déclin du fait de leur coût important et du faible degré d'accessibilité de ces démarches et de leurs résultats pour les décideurs. Dans les années 1980, vont ensuite dominer les approches de type *value for money* visant une meilleure allocation des ressources mais dans une perspective principalement micro-économique, déconnectée du contexte et sans prise en compte des effets sociaux et qualitatifs autre que budgétaires (Mairate, 2007).

Il faut donc attendre le milieu des années 1990 pour que soit développée la pratique évaluative dans le domaine du développement régional. L'évaluation est introduite dans le cadre de la politique régionale communautaire avec le « Paquet Delors » II adopté lors du Conseil européen d'Edimbourg et qui couvre la période de programmation 1994-1999. Le poids des fonds structurels est doublé en passant au maximum historique de 0,46 % du PIB communautaire et le cadre évaluatif est précisé. Le processus est ensuite renforcé à l'occasion de la démarche « Agenda 2000 » préparant l'élargissement de l'Union européenne et le renforcement des procédures de suivi et d'évaluation. Le cadre évaluatif est quant à lui formalisé à travers un règlement de 1999 qui rend obligatoire l'évaluation sous peine d'entraîner un refus de concours financiers à un Etat-membre. Ce règlement dispose en effet que l'attribution de la « réserve de performance », à savoir les bonus financiers s'ajoutant aux financements contractuels (soit 16 Mds€ pour la France en 2003), est conditionnée à la réalisation effective des évaluations.

Quatre principes sont formalisés pour mettre en œuvre les activités d'évaluation au niveau régional :

- la programmation multi-annuelle qui confère stabilité et cohérence aux politiques,
- l'additionnalité budgétaire des financements européens qui ne doivent pas se substituer aux fonds nationaux,
- la concentration des crédits européens dans des zones identifiées comme prioritaires,
- et enfin la conduite de politiques partenariales associant les différents niveaux d'action au sein des Etats-membres tant dans la conception des programmes que dans leur mise en œuvre.

En pratique, l'évaluation dans les programmes communautaires se caractérise par trois spécificités. Elle porte tout d'abord sur des programmes complexes impliquant une multitude d'acteurs avec des niveaux de responsabilité variés et des objectifs interdépendants. Elle se singularise ensuite par la dimension décentralisée de sa gestion, au niveau régional ou local, voire déléguée à des organismes en dehors du cadre des administrations publiques. Le principe de subsidiarité constitue aussi un principe essentiel du déploiement des évaluations dans la mesure où ces dernières visent avant tout à démontrer la valeur ajoutée de l'intervention communautaire. Enfin, les programmes communautaires impriment une temporalité particulière aux exercices d'évaluation : planifiés selon des cycles de 5 à 7 ans afin de favoriser des effets d'apprentissage ces programmes permettent aux évaluations d'apporter leur pleine mesure en termes d'accumulation de connaissances.

Sur le plan méthodologique, le cadre communautaire d'évaluation impose l'organisation des évaluations selon trois phases :

- La première phase, consacrée à l'évaluation *ex-ante*, est obligatoire et vise tout à la fois à apprécier le positionnement du programme par rapport au diagnostic du territoire, se prononcer sur la pertinence de ses objectifs au regard des besoins, dresser un bilan socio-économiques des zones éligibles et définir les indicateurs de suivi nécessaires à mettre en place. Les études d'impact réalisées dans de nombreux programmes s'inscrivent dans ce type d'évaluation *ex-ante*.

- La deuxième phase consiste en une évaluation à mi-parcours ou intermédiaire. Elle comprend la réalisation d'un bilan d'avancement du programme, une analyse des conditions de sa mise en œuvre et une appréciation des premières réalisations. L'ensemble de ces analyses est destiné à apporter les ajustements nécessaires en cours de route.
- L'évaluation finale intervient quant à elle lors de la troisième phase, en fin d'exécution du programme. Elle permet avant tout de rendre compte de l'efficacité des actions et de tirer les enseignements d'ensemble à prendre en compte au moment de préparer le nouveau cadre de programmation.

La Commission Européenne a joué un rôle moteur dans la diffusion de l'évaluation à travers quatre démarches complémentaires (Mairate, 2007, p.31). L'initiative *Sound and efficient Management* lancée en 1995 a d'abord permis de généraliser la création d'unités spécialisées d'évaluation au sein de la commission européenne³⁸. La Commission européenne a également été pionnière dans la construction de standards d'évaluation. L'ensemble des manuels d'évaluation réunis dans la collection MEANS réalisée au début des années 1990 matérialise ce processus de production de règles et de bonnes pratiques méthodologiques. Une attention particulière a également été portée par les services de la Commission dans l'assistance permanente aux Etats-membres tant pour l'accompagnement dans la mise en place de référentiels d'évaluation que pour la résolution de problèmes concrets posés aux gestionnaires nationaux par exemple pour le suivi d'indicateurs ou la définition des outils d'évaluation. Enfin, dès 1992, la Commission européenne a été à l'origine d'un réseau rassemblant des évaluateurs des différents Etats-membres dans le but de partager les expériences nationales en matière d'évaluation.

Dans la mesure où elle appelle la mobilisation d'importants financements européens, la mise en œuvre des Contrats de Plan Etat Région (CPER) au cours de la décennie 90 a été fortement marquée par ce cadrage communautaire. Il a été par la suite étendu à d'autres procédures contractuelles mises en œuvre entre l'Etat et les collectivités. Issus de la planification, les CPER sont un instrument de contractualisation entre l'Etat

³⁸ Commission européenne, *Concrete steps towards best practice across the commission*, 8th May 1996, SEC 96/651 final.

et les Régions autour de projets cofinancés dans le domaine de l'aménagement et du développement régional. Bien qu'existant depuis la loi de 1982 portant réforme de la planification (loi n°82-653 du 29 juillet 1982), les CPER sont progressivement devenus des instruments privilégiés de la négociation entre l'Etat et les collectivités territoriales dans la définition des objectifs de l'action publique et l'allocation des ressources. Le recours à la démarche contractuelle a en effet contribué à modifier l'équilibre de la décision et à accompagner le passage d'une action centralisée dans laquelle le contrôle bureaucratique, principalement du respect des formes juridico-administrative occupait une place centrale, à une action par projets s'appuyant classiquement sur la détermination d'objectifs préalables, observables et mesurables.

L'obligation d'évaluer les politiques publiques au sein des CPER résulte de la traduction dans le droit interne des procédures communautaires (règlement CEE du Conseil du 20 juillet 1993 rendant obligatoire l'évaluation des opérations financées par avec l'aides des fonds européens). La circulaire du 9 décembre 1993 en précise les principes et l'organisation pour une application effective à la troisième génération des CPER couvrant la période 1994-1998. Les crédits nécessaires aux évaluations sont fixés à hauteur de 6 / 10 000 ème du volume des crédits, soit à l'époque entre 1,5 et 3 MF par région.

Au début des années 2000, les bilans et retours d'expérience réalisés à la fin de la première vague d'évaluation des CPER sur la période 1994-1999 faisaient état d'un démarrage difficile des évaluations. Les crédits dédiés aux évaluations n'avaient été consommés que partiellement et aléatoirement. Ces évaluations ont révélé également des dysfonctionnements liés à la multiplication des acteurs et à l'hétérogénéité des systèmes d'information. Le rapport remis en 2000 par le Sénat partageait ce bilan critique en soulignant la portée limitée de l'évaluation dans ces premières procédures contractuelles. Il mettait aussi en évidence la confusion des objectifs résultant du mode de décision partenarial ainsi que l'appauvrissement de l'évaluation lié au fractionnement des objets à évaluer. Une enquête menée dans quatre régions française dont la Bretagne, avait aussi conclu à un impact faible des évaluations sur les pratiques. Selon cette étude, les évaluations affectaient peu les politiques budgétaires et ne nourrissaient pas de réflexions approfondies sur les logiques financières. Parfois, les résultats avaient été intégrés aux documents budgétaires mais cette pratique était loin

d'être systématique. Enfin, les évaluations ne se traduisaient pas davantage par une réorientation des partenariats. Pour autant, ces différentes analyses montraient que l'exercice n'avait pas été vain et que l'évaluation avait permis des effets d'apprentissage tant du point de vue des coopérations entre acteurs que sous l'angle d'une meilleure connaissance des actions mises en œuvre.

L'engagement des collectivités territoriales dans l'évaluation : quelques exemples d'évaluations régionales et locales

Des effets d'apprentissages analogues peuvent être observés au niveau des initiatives propres menées par les collectivités territoriales où le développement de l'évaluation a été porté par des dynamiques davantage volontaristes que contraignantes, mais également plus expérimentales et aléatoires. L'institutionnalisation de l'évaluation au sein des collectivités locales a, en effet, été développée lentement et irrégulièrement, d'abord par imitation des pratiques nationales puis à l'occasion de leur reconnaissance par l'Etat dans le cadre des Contrats de Plan (Hauraist, 1992). Des structures d'évaluation ont été mises en place d'abord par des collectivités de niveau régional, le plus souvent dans un souci d'autonomisation et de légitimation de cet échelon, avant d'être suivis par des conseils généraux pour les compétences sociales des départements et par les collectivités infra-territoriales (communes et intercommunalités) dans le cadre notamment des évaluations de la politiques de la ville. Quelques exemples parmi les plus marquants peuvent être rappelés ici pour ces différents niveaux.

- Région Rhône-Alpes : au niveau de la Région Rhône-Alpes, la décision de créer à titre expérimental un dispositif d'évaluation des politiques régionales remonte à 1991, et a été reconduit en 1995 et 1996 (Auvolat, 1996). L'originalité du dispositif rhônalpin tient à ce que l'assemblée délibérante constitue l'instance d'évaluation, c'est-à-dire que la commande et la conduite de l'évaluation sont confondues, à la différence du modèle habituel où la séparation est pratiquée au nom de l'indépendance. Si les études d'évaluation sont réalisées par un prestataire extérieur et que deux comités, le comité scientifique et le comité de pilotage, veillent à la rigueur méthodologique et au suivi des travaux, le monopole des propositions d'évaluation revient à l'exécutif. Chaque année, des évaluations sont ainsi décidées et s'ajoutent aux

évaluations obligatoires du contrat de plan. Entre 1991 et 1996, les évaluations réalisées ont essentiellement porté sur les compétences propres de la Région telles que l'innovation et les transferts de technologie en faveur des PMI (1992), la formation professionnelle en alternance (1994), les programmes intégrés de développement agricole (1994), la recherche (1995), le programme spécifique de soutien scolaire, d'orientation et d'ouverture culturelle des élèves (1996) et l'habitat (1996). A travers cette sélection de politiques novatrices décidées unilatéralement, l'ambition de l'exécutif régional consistait à démontrer l'existence et la valeur des politiques régionales. Cette recherche d'auto-légitimation a ainsi conduit les responsables de l'évaluation au Conseil régional à recourir à des approches sociologiques et qualitatives afin de rendre compte le plus concrètement possible de l'intervention régionale et de ses effets. Dans ses cahiers des charges la Région appelait ainsi les évaluateurs à apprécier la mise en œuvre des actions, en particulier la pertinence des structures et des procédures imaginées par la région. Il s'agissait également d'analyser plus globalement la pertinence et la cohérence de la politique régionale, en particulier dans son articulation avec d'autres actions conduites par d'autres acteurs et d'insister à travers l'évaluation sur les effets structurants des actions, qu'il s'agisse d'effets leviers ou d'effets organisations (Warin, 1996).

- Région Bretagne : le développement de l'évaluation par le Conseil régional de Bretagne au début des années 1990 procède d'une dynamique différente en raison du contexte historique et culturel comme du point de vue des finalités recherchées. Il est le fruit de la conjonction de plusieurs volontés modernisatrices et notamment celles des cadres dirigeants de la Région cherchant à sensibiliser les services au « carré magique » (préparation – planification – exécution – contrôle/évaluation) alliées à celles d'un préfet réputé réformateur et du Président de la Chambre Régional des Comptes s'inscrivant délibérément dans le cadre des recommandations gouvernementales de la directive Rocard de « renouveau des services publics » (Fontaine, 1996). L'évaluation y est clairement définie dans une optique d'aide à la décision. Le rôle du Président du Conseil régional et des élus est faible et le dispositif évaluatif breton épouse la tradition historique de partenariat entre l'Etat et la Région Bretagne, co-gestionnaires de

l'action publique régionale. Le dispositif se compose de 4 instances : le comité régional de l'évaluation (CRE), le comité scientifique (CS), un groupe de suivi et les chargés d'évaluation. Le Comité scientifique réunit avant tout des contrôleurs de gestion et des économistes et statisticiens, traduisant ainsi l'approche « technico-économique » du modèle évaluatif breton, qui privilégie le « monitoring » et marginalise les interrogations plus larges sur le bien fondé des politiques ou sur l'utilité des actions. La définition retenue s'inscrit dans la veine du rapport Deleau de 1986 puisqu'il s'agit de « reconnaître et de mesurer les effets propres d'une politique ». A travers la construction d'une douzaine d'indicateurs de programme, à la fois « représentatifs et opérationnels » (Baslé, Pelé, 1994), les promoteurs de l'évaluation attendent le développement d'un « esprit d'économicité » dans les administrations en charge des politiques et ambitionnent rien de moins que « l'objectivation scientifique par l'évaluation » (Fontaine, 1996). Impliqué dans le comité scientifique en tant que chercheur, Joseph Fontaine conclut de l'expérience bretonne que, 5 ans après son lancement, la démarche d'évaluation n'a pas modifié la préparation et le suivi des politiques pour ce qui concerne l'exécutif et l'assemblée de la Région. Elle n'a pas représenté une opportunité pour modifier l'importance des formes de la consultation, de la délibération et de négociation de la production des politiques publiques. Les Conseillers régionaux n'ont été que très épisodiquement informés des résultats des évaluations, et sans ce que cette situation ne suscite de réactions de leur part. Selon Fontaine, on retrouve dans l'évaluation telle qu'elle est déployée dans le cas breton ce qui est révélé de manière récurrente par les analyses sur le management public à savoir le primat des impératifs de gestion interne, la volonté de contrôle des centres d'expertises indépendants et l'absence de prise en compte de la demande sociale.

Outre ces deux initiatives pionnières, plusieurs Conseils régionaux ont développé depuis des activités d'évaluation, à des degrés certes plus ou moins formalisés³⁹ mais qui font des Régions des commanditaires de premier plan en ce qui concerne les évaluations, à travers les CPER mais aussi à travers bon nombre de politiques régionales. Les exemples bretons et

³⁹ Notamment pour les Régions Pays de la Loire, Lorraine, Limousin ou Nord-Pas-de-Calais.

rhônalpins, largement commentés dans la littérature, restent néanmoins d'actualité pour analyser la mise en œuvre de l'évaluation au niveau régional. Ils illustrent à la fois le poids des contextes régionaux et des configurations d'acteurs dans la mise en place de l'évaluation et le rôle de l'évaluation comme élément de légitimation de l'action publique territoriale. Dans les deux cas, l'évaluation est d'abord mobilisée pour ses usages stratégiques avant de remplir d'autres finalités : cognitives, managériales ou démocratiques. Elle est en effet considérée par les acteurs qui s'en saisissent comme un levier pour renforcer leurs positions et améliorer leur action. Elle est également perçue comme une source forte d'autonomisation, que ce soit à travers l'autonomie d'expertise dans le cas de la Bretagne ou à par le biais de l'affirmation du rôle entrepreneurial du conseil régional dans le cas de rhônalpin (Warin, 1996).

Au niveau infra-régional, l'Hérault fait partie des départements ayant mis sur pied, dès les années 1990, leur propre dispositif d'évaluation. L'approche développée par le Conseil général de l'Hérault a été inspirée du dispositif interministériel, composée d'une instance décisionnelle et d'un comité scientifique, et dotée de crédits spécifiques (Laget, Benamour 1996). La création d'un dispositif départemental d'évaluation procède d'une volonté politique de transparence et de maîtrise des actions du Conseil général, dans un contexte de décentralisation et de crise économique se répercutant sur la gestion des collectivités locales.

Le choix des évaluations revient à une instance issue de l'exécutif départemental, le comité départemental d'évaluation, présidé par le Président du Conseil général et composé de 7 Vice-présidents délégués de l'exécutif. Un conseil scientifique rassemblant 7 personnalités indépendantes est chargé du suivi méthodologique et donne un avis sur les documents établis dans le cadre de l'évaluation. Ce dispositif peut également être saisi d'une demande d'évaluation par des communes ou des autorités administratives. Plusieurs évaluations ont été ainsi lancées depuis 1990, sur la politique de l'emprunt, du patrimoine, du développement local ou sur la politique de l'eau. Cette dernière évaluation, réalisée en 1992 a notamment eu pour effet de confirmer le champ d'une politique proprement départementale, même si la collectivité est tributaire d'autres maîtres d'ouvrage. Elle a en outre conduit à la mise en place de trois schémas prospectifs : alimentation en eau potable, assainissement, valorisation des boues d'épuration.

Après trois ans de pratique, le bilan dressé de l'expérience Héraultaise faisait ressortir plusieurs difficultés parmi lesquelles la confusion persistante entre l'évaluation et les démarches de contrôle et d'audit, le manque d'appropriation voire les résistances de la part des élus de l'assemblée départementale, mais aussi le faible impact sur le débat public des rapports d'évaluation pourtant diffusés dans le but explicite de nourrir le débat démocratique local. S'agissant des apports, il était noté que l'évaluation avait rempli pour l'essentiel des fonctions de connaissance en permettant de mieux formaliser certaines politiques publiques jusque là considérées comme « implicites ». Ce qui a par la suite facilité la prise de décision des élus et de clarifier le rôle et les interventions des différents acteurs institutionnels dans la définition et la mise en œuvre des politiques complexes.

Au niveau local, outre le fait que, contrairement aux Région, l'évaluation n'a pas été imposée de l'extérieur par les dispositifs contractuels, les démarches évaluatives ont longtemps été freinées par les caractéristiques propres à la gestion locale. Parmi celles-ci on peut rappeler le fait, tout d'abord, que les collectivités locales sont faiblement pourvues en termes de capacité d'analyse et d'audit, ne disposant pas en propre de corps d'inspection, sauf exceptions (inspection générale de la ville de Paris). Elles ont également longtemps pâti d'un développement insuffisant des outils de pilotage et des fonctions relatives au contrôle de gestion. En outre, la formulation des politiques locales n'est pas toujours explicite : les délibérations ne sont pas précédées d'exposés de motifs, comme le sont les textes législatifs et réglementaires. Enfin, le jeu partenarial au local rend encore plus ardue la tâche de définir des objectifs clairs et des règles du jeu uniformes. L'incapacité d'isoler localement les effets parasites des autres politiques fait donc que l'on évalue le plus souvent des politiques plurielles territorialisées quand il ne s'agit pas d'« une espèce de mélange de politiques qui s'inscrivent sur un territoire » (Thoenig, 1998, p. 313-329).

En 2002, une enquête menée par le cabinet Bernard Bruhnes Consultants sur les pratiques managériales dans les collectivités locales, en partenariat avec l'association des administrateurs territoriaux de France⁴⁰, soulignait la

⁴⁰ Bernard Bruhnes Consultant (en partenariat avec l'Association des Administrateurs Territoriaux de France), « Vingt ans de décentralisation. Où en est aujourd'hui le management public territorial ? » *Enquête sur les pratiques managériales*

timidité des démarches évaluatives au niveau local. Sur les 46 collectivités ayant répondu à l'enquête, 52 % déclaraient procéder à l'évaluation de certaines de leurs politiques. Les plus concernées étant les politiques contractuelles, ainsi que les politiques consacrées à l'habitat, la petite enfance et l'insertion. Pour ces collectivités concernées, l'évaluation est réalisée le plus souvent en partenariat avec d'autres collectivités ou services de l'Etat. Enfin, dans seulement 29 % des cas, les collectivités indiquaient disposer d'un service ou d'une direction de l'évaluation.

Encart n° 14 : Enquête de 2002 sur les démarches d'évaluation dans les collectivités territoriales

Les démarches d'évaluation restent encore timides

52 % des collectivités interrogées déclarent procéder à l'évaluation de certaines de leurs politiques :

- les politiques les plus fréquemment concernées sont les politiques contractuelles (contrat de Plan État/Région, politique de la Ville, PLIE, contrats locaux de sécurité...), ainsi que les politiques d'habitat, petite enfance, insertion ;
- 85 % des collectivités procédant à des évaluations affirment conduire ces démarches en partenariat.

Il n'existe de dispositif permanent d'évaluation que dans 16 % des collectivités consultées, et 29 % des collectivités interrogées déclarent disposer d'un service ou d'une direction de l'évaluation.

Les évaluations réalisées en interne font le plus souvent appel à des outils d'usage courant (tableaux de bord...) le plus souvent auto-administrés par les services opérationnels porteurs de l'action ou du projet évalué.

Les démarches d'évaluation paraissent plus souvent imposées (notamment par les dispositifs contractuels) que volontaires.

Pour les collectivités qui les pratiquent, les démarches d'évaluation sont effectuées :

- chemin faisant pour 54 % des collectivités concernées ;
- ex-ante pour 16 % des collectivités concernées ;
- ex-post pour 29 % des collectivités concernées.

Les évaluations mises en œuvre cherchent à mesurer :

- l'efficacité (rapport entre les résultats observés et les résultats visés) : 36 % ;
- l'efficience (rapports entre les moyens engagés et les résultats) : 9 % ;
- l'impact (ensemble des effets produits) : 55 %.

Plus récemment, une enquête demandée par l'Institut des villes, permet de mesurer les progrès de l'évaluation dans les villes de plus de 50 000 habitants. Parmi les 48 villes ayant répondu (sur les 118 concernées), 85 % d'entre-elles déclarent l'évaluation comme connue ou familière. Mais elles sont souvent confondues avec l'audit ou le contrôle. Les pratiques évaluatives volontaires ou obligatoires existent dans les deux tiers des villes, depuis plus de cinq ans pour 30 % d'entre-elles. Ces pratiques ont émergé à la faveur des obligations contractuelles (32 %) et/ou du fait de la volonté des élus locaux (27,5%), mais aussi sous la pression des collectivités partenaires et dans un contexte de « métropolisation » de l'action locale. Il reste que

l'évaluation sur des actions en cours est privilégiée et que dans 90 % des cas les travaux sont jugés utiles.

Dans ce contexte, la politique de la ville fait office d'exception. À défaut d'avoir suscité des évaluations dignes de cette appellation, l'incitation nationale à l'évaluation locale des dispositifs successifs de la politique de la ville (conventions « habitat et vie sociale », développement social des quartiers...) a contribué à l'institutionnalisation progressive de cette pratique (encart n°15). S'appuyant sur une enveloppe financière conséquente et sur la diffusion d'un guide méthodologique pour l'évaluation édité par la Délégation interministérielle à la ville (DIV), les vagues répétées d'évaluation des contrats de ville ont indéniablement contribué au développement de la pratique évaluative dans les administrations déconcentrées et les collectivités. Entre 1994 et 1999, plus de la moitié des contrats de ville avait ainsi fait l'objet d'une évaluation intermédiaire ou finale (Epstein, 2010). Bénéficiant le plus souvent du concours d'organismes privés ayant structuré leur offre d'expertise (cabinets d'audit, bureaux d'études, experts indépendants...), ces évaluations ont souvent eu une fonction pédagogique (Spencehauer et Warin, 2000), suscitant des dynamiques d'apprentissage collectif qui ont facilité la mise en œuvre d'une politique impliquant de multiples institutions. Dans certains sites, ont été mis en place des tableaux de bord agrégeant les données statistiques et de gestion disponibles, palliant les carences d'un appareil statistique national qui est structuré pour répondre aux besoins de connaissances de l'État et non à ceux des pouvoirs locaux (Desrosières, 1993). De manière plus frappante, l'évaluation est parvenue dans certains cas à faire émerger des scènes de débat local sur les enjeux, les objectifs et les réalisations de la politique de la ville, qui ont permis d'infléchir la politique menée (Kirszbaum, Epstein, 2005).

En dépit de ces succès, la politique de la ville a connu ces dernières années un glissement progressif vers des démarches d'observation au détriment d'activités plus riches d'évaluation. S'appuyant sur un rapport de la Cour des Comptes de 2002 qui concluait à une politique de la ville impossible à évaluer du fait de ses « tares conceptuelles », en particulier le flou de ses objectifs et l'attribution de responsabilité dans la mise en œuvre. La réforme introduite par la loi d'orientation et de programmation pour la politique de la ville et de la rénovation urbaine du 1^{er} août 2003 (dite « Loi Borloo ») a été en grande partie justifiée par un impératif d'évaluation. Dans le cadre de

cette réforme, plusieurs efforts ont été entrepris pour rendre évaluable la politique de la ville à travers la redéfinition des enjeux et des objectifs de la politique publique en fonction des instruments de mesure des résultats, la production de batteries d'indicateurs quantitatifs ou encore la création d'une Agence nationale de la rénovation urbaine (ANRU) supposée rendre plus lisible le pilotage institutionnel des actions. Mais ces améliorations ont surtout permis de faire progresser la connaissance plus qu'elles n'ont rendu possible la mesure de l'efficacité des différents programmes et la formulation d'un jugement sur l'efficacité d'ensemble de la politique. Si les différents outils statistiques ont permis des avancées importantes pour faire progresser la connaissance des quartiers concernés par la politique et des moyens affectés, ils n'ont pas fourni les conditions suffisantes pour que les programmes fassent effectivement l'objet d'évaluations. En particulier, les responsables des programmes de rénovation urbaine ont fait preuve d'une grande prudence, s'interdisant de prolonger leurs analyses statistiques par un jugement sur la valeur des différentes mesures de la politique de la ville. Le Comité d'évaluation et de suivi de l'ANRU a été plus audacieux, publiant plusieurs avis et rapports critiques sur les moyens dévolus à la rénovation urbaine ou sur les conditions de pilotage du programme mais ces derniers n'ont apporté aucun élément sur son efficacité et son impact (Epstein 2010). On notera toutefois avec Renaud Epstein que ce déficit d'évaluation n'a pas empêché les responsables de la politique de la ville de célébrer la réussite du Programme national de rénovation urbaine (PNRU) et les parlementaires de voter à trois reprises la prolongation de cet ambitieux programme qui prévoyait la démolition et la reconstruction de 250 000 logements sociaux, pour un coût total estimé à 42 milliards d'euros.

Encart n° 15 : L'évaluation de la politique de la ville en France

Dès les années 1980, la politique de la ville a en effet servi de terrain privilégié pour l'expérimentation des premières démarches d'évaluation en France (Ion, Peroni, Roux, 1996) puis de « cheval de Troie pour la diffusion de cette pratique dans les appareils administratifs au cours de la décennie suivante » (R. Epstein, 2010). Mais ces démarches entreprises d'abord dans le domaine de la politique de développement social des quartiers (DSQ) puis prolongées dans le cadre des travaux du Commissariat Général du Plan et, plus tard, du Comité national d'évaluation créée sous l'égide du Conseil National des Villes, ont surtout permis d'éclairer les procédures d'évaluation au niveau national, en inspirant notamment le dispositif institutionnel CIME-CNE mis en place à partir de 1990. Au niveau des collectivités locales l'évaluation est surtout apparue comme un levier de mobilisation des acteurs locaux dans les débats de proximité sans que ne soit pensé un dispositif de mise en place des capacités d'évaluation des villes pour accompagner ce renvoi de l'évaluation au niveau local (Epstein, 2007).

L'analyse des différentes évaluations des contrats de ville sur la période 2000-2006 concluait ainsi que ces évaluations locales reproduisaient les travers méthodologiques observées quelques années plus tôt. De plus, l'analyse des procédures d'évaluation analysées montraient que ces dernières étaient souvent contrariées par de fortes tensions entre les différentes parties prenantes ne parvenant pas à s'accorder sur la nature et les finalités mêmes de la politique de la ville. Pour certains, il s'agissait d'une politique « substitutive » venant traiter les territoires et les populations délaissés par les politiques sectorielles. Pour les autres, elle incarnait une politique « incitative » visant à agréger et articuler les différentes politiques sectorielles afin qu'elles parviennent à prévenir et résorber les situations d'exclusion. Faute d'accord sur ce point, le passage des connaissances produites à un jugement partagé sur la valeur de la politique de la ville était impossible (Epstein, 2010).

2.4. Autre acteurs, autres rattachements

Deux autres solutions organisationnelles : l'INSEE et une autorité administrative indépendante

A deux périodes très différentes deux autres solutions organisationnelles ont été envisagées en France pour accueillir les activités d'évaluation. L'INSEE (années 1970) et une Autorité administrative indépendante de l'évaluation (années 2000). Elles sont de nature profondément différentes,

mais se rejoignent sur trois points indicatifs des enjeux récurrents en ce domaine. Tout d'abord l'une et l'autre ont été envisagées comme une solution à la délicate distance souhaitable avec les autorités politiques. Comment établir du lien sans subir de la dépendance ? Comment participer aux choix d'action publique majeurs sans céder au besoin de légitimité à court terme de l'action gouvernementale ? Ensuite, elles ont été considérées comme ayant des statuts institutionnels suffisamment forts pour être capables de limiter l'interférence avec les besoins à court terme des décideurs. Dans l'un et l'autre cas, c'est la capacité à développer une expertise technique de très haut niveau qui justifiait aussi le choix. Enfin, dans les deux cas, c'est aussi la possibilité d'accès et si besoin de production des informations (source indispensable à l'évaluation et dans laquelle s'épuisent beaucoup d'instances) qui était mis en évidence.

***L'INSEE et le projet d'un « marché » de l'information et de l'étude »
(fin 60 – début des années 70)***

Dans sa recherche très approfondie sur les origines et le parcours chaotique de l'évaluation dans l'administration française, V. Spenlehauer (1998) analyse un épisode négligé, celui du projet qui a émergé au sein de l'INSEE de s'impliquer dans ce type de production de connaissances. Ce que l'auteur résume dans une formule synthétique « le rôle et la place paradoxaux de l'institution INSEE (...) dans le champ de la pratique de l'évaluation des politiques publiques. En effet, ce rôle et cette place se caractérisent à la fois par la quasi absence de l'organisme administratif et par une remarquable présence des administrateurs de l'INSEE » (p. 158).

Fin 1966, Claude Gruson qui a fortement contribué au développement de l'organisme quitte la direction de l'INSEE. Il est remplacé par Jean Ripert qui vient du Commissariat général du Plan (CGP). Celui-ci a suivi les évolutions du CGP et l'emprise de la logique économique sur les activités de planification. Il souhaite que dans les choix d'action publique les questions sociales et économiques soient davantage articulées. Cette préoccupation rencontre celle d'une partie montante des agents de haut niveau de l'INSEE « qui supportent de plus en plus difficilement d'être de simples techniciens statistico-économiques au service du CGP qui fait en quelque sorte office d'écran entre eux et les utilisateurs de leurs travaux ... (il apparaît) à un nombre croissant de cadres de l'INSEE comme un débouché à la fois trop

étroit et trop grossier pour leur production » (p.160). L'organisme va alors développer une « nouvelle approche de la question sociale » afin de rénover le profil de l'INSEE (en prenant des distances avec le modèle de la Comptabilité publique traditionnelle) et de renforcer son autonomie tant institutionnelle qu'intellectuelle. Un colloque est organisé en juin 1967 sur le rôle de l'information statistique dans la société (entendue au sens large). « Le projet intellectuel que nourrissent les jeunes turcs de l'INSEE consiste à pluraliser voire à démocratiser l'usage de l'information économique et sociale produite par l'Institut » (p. 162). A l'inverse de la logique centralisée et dirigiste du CGP dont l'action est jugée déresponsabilisante, l'INSEE veut inciter à la création d'observatoires économiques régionaux (en coopération avec la DATAR) et de services d'études dans tous les établissements publics (administrations, entreprises nationales, collectivités locales) qui utiliseraient pour leurs actions les données scientifiquement validées de l'INSEE.

Deux facteurs conjoncturels vont accélérer l'évolution de l'INSEE. D'une part, les mouvements politiques de mai 68 relayés dans l'organisme par une nouvelle génération d'économistes-statisticiens, d'autre part l'intérêt croissant porté par le CGP (par l'intermédiaire de J. Delors) à l'élaboration d'indicateurs sociaux. Dans ce contexte, certains chercheurs de l'INSEE envisagent de développer une « évaluation *ex-post* objectiviste des programmes publics à composante sociale » (p. 184). Le projet ne fait pas l'unanimité, mais il est étoffé et débouche sur un texte d'avril 1972 (en fait, il s'agit d'une note n°14c des collections de l'INSEE) au sein du groupe « statistiques sociales » qui propose de combiner production de connaissances et aide à la décision. Contrairement aux analyses faites au CGP sur les indicateurs sociaux qui supposent « qu'il suffit de produire de l'information sur la société pour que les décideurs modifient rationnellement et comme par enchantement l'action publique » (p. 185), l'argumentation de l'INSEE considère qu'un « système de statistiques sociales devra aussi englober le fonctionnement de l'action publique dans son objet » dans la mesure où il fait partie du fonctionnement social global. V. Spenlehauer montre que l'analyse faite alors à l'INSEE dépasse le modèle de référence étasunien de « *social report* » et s'inspire davantage du modèle « *social trends* » anglais en se proposant de prendre en compte les systèmes de valeur de l'action publique toujours occultés, mais aussi le flou et la complexité (voire les contradictions) dans les objectifs des programmes publics analysés. Ce raisonnement qui intègre la question des systèmes de valeur des choix publics et les ambiguïtés de

leurs objectifs montre à quel point la problématique a évolué, partant d'une approche technicienne rationaliste à une approche politique. Le document de l'INSEE est d'ailleurs explicite sur ce point lorsqu'il indique que « la référence aux objectifs (de l'action publique) n'implique pas qu'ils fassent l'objet d'un quelconque consensus ou que les objectifs soient réellement poursuivis : elle signifie seulement que leur formulation et leur degré de priorité ont une signification politique » (p. 44 de la note n°14c).

Mais faute de soutien en interne, cette approche « *policy oriented* » développée par une partie de la nouvelle génération de l'INSEE ne prospère pas au sein de l'institution : « Ceci s'explique par le doute profond qu'éprouve le duo responsable de la division des programmes quant à la possibilité de rationaliser véritablement les processus de décision politico-administrative » (p.191).

Une autorité administrative indépendante ?

Face à l'éparpillement administratif des pratiques actuelles et face aux limites de la LOLF que peut-on proposer pour donner un sens renouvelé aux activités d'évaluation ? Le rapport parlementaire Bourdin et al. (2004) voit une solution dans la création d'une nouvelle institution la Haute autorité de l'évaluation (HAE) qui superviserait un ensemble de structures. La Haute autorité aurait pour tâche principale la définition et le contrôle de « bonnes pratiques de l'évaluation » par la validation de méthodes et leur diffusion auprès des acteurs concernés (via des actions d'information et de formation). Elle serait aussi un superviseur des actions évaluatives menées ; enfin, elle définirait et garantirait les règles déontologiques de fonctionnement des instances.

Une seconde mission complèterait de façon plus originale la première, la HAE aurait aussi en charge de « favoriser l'efficacité de l'évaluation ». Dans ce sens il lui appartiendrait tout d'abord de veiller au bon déroulement des évaluations tout au long de leur accomplissement (en réglant les tensions entre commanditaire et acteurs de l'évaluation, en débloquent l'accès à des sources d'information, etc.). Ensuite, elle serait en charge de la lisibilité des travaux d'évaluation et du suivi de leurs effets. C'est le caractère très délicat et potentiellement conflictuel de ces missions à accomplir dans le contexte politico-administratif français déjà évoqué, qui impose selon les rapporteurs

le recours à une autorité dotée de la plus grande indépendance. Les opérations d'évaluation elles-mêmes seraient menées à partir de Commissions de l'évaluation créées par grands secteurs de l'action publique et organisées sur un mode pluraliste (sur un modèle proche de ce qu'était le Conseil national de l'évaluation). Celles-ci sélectionneraient les demandes et reformuleraient au besoin les attentes du commanditaire. Elles formeraient des instances *ad hoc* en fonction des projets retenus.

Les parlementaires entendent ainsi doter la Haute autorité d'une légitimité particulière qui lui permette de résister aux pressions politiques et aux résistances corporatistes de la haute administration et des grands corps. Ce projet s'inscrit en effet en continuité avec les attributs classiquement attribués aux Autorités administratives indépendantes, en particulier celles relevées par le Conseil d'Etat en 2001 qui voit en elle des instances dotées d'une crédibilité particulière (Gentot, 1991). Cette dernière se déclinant en termes d'impartialité, de professionnalisme et de souci d'efficacité. Le Conseil en fait aussi un « véritable stimulant de l'action publique » en raison des médiations que les AAI opèrent entre acteurs publics cloisonnés et entre eux et la société civile. Mais prendre ainsi le problème par le haut en ajoutant à l'édifice déjà complexe une couche supplémentaire, est-ce le meilleur moyen de traiter des pesanteurs institutionnelles qui resteraient inchangées ? Sur la base d'un travail de comparaison internationale S. Jacob est très réservé sur l'efficacité de ces organismes centralisés et unificateurs. Il met plutôt l'accent sur le rôle moteur de réformateurs administratifs et d'entrepreneurs politiques dans la diffusion et l'appropriation des pratiques évaluatives.

Deux points de discussion peuvent être introduits sur cette proposition de création d'une AAI en matière d'évaluation. Tout d'abord, on comprend bien dans ce projet le souci d'échapper d'un côté aux intrusions et instrumentalisation politiques, et d'un autre côté à l'emprise technocratique des grands corps et à leur prétention au monopole de l'expertise. Mais comme l'ont montré différents travaux sur les pratiques d'évaluation (Delley, 2006), la tenue à distance du politique ne peut être une fin en soi. Au contraire, la validité des commandes en amont et la réception des conclusions évaluatives sont inséparables de la qualité des échanges entre évaluateurs et responsables des *policies* et du *politics*. Ce sont ces relations qu'il faut pouvoir travailler et faire évoluer pour sortir l'évaluation de son rôle mineur. J. Leca n'a cessé de relever le caractère central de cette articulation incontournable (Leca 1993 ; Duran 1999 ; Gaxie Laborier 2003).

Ensuite, au vu de ce que les sciences sociales ont apporté depuis cinquante ans à l'analyse des processus de changement dans les institutions, ne faudrait-il pas dans un premier temps privilégier l'apprentissage des acteurs de l'action publique avant de prétendre modifier l'institutionnalisation ? Tous les observateurs, le disent et le répètent (et S. Jacob avec une netteté particulière) ce qui fait défaut en France ce sont les prémisses cognitives de l'action évaluative et l'absence de cette dimension dans les cultures professionnelles des acteurs publics (Chanut, 2002). La réalisation de cette première suppose l'insertion d'une culture d'évaluation dans la formation ; elle peut aussi être accomplie (ou renforcée) par la découverte expérimentale des enjeux et problèmes de base de l'évaluation lors d'opérations internes mobilisatrices. Au sein de tous les départements ministériels des opérations d'évaluation pourraient ainsi être menées à l'occasion des premiers suivis de la LOLF sur la base d'une méthodologie commune. Par exemple, des actions développées par des services de première ligne pourraient toutes s'efforcer de répondre à la question suivante : dans quelle mesure la nouvelle nomenclature budgétaire permet-elle de rendre comptes des actions menées ? Il s'agirait d'explicitier la réalité dont essayent de rendre compte les indicateurs choisis. Et de répondre aux questions de base : que mesure-t-on ? Comment cela est-il effectué ? Quelles significations sont ainsi produites ? Ce serait autant un moyen de mettre à l'épreuve les catégorisations et les indicateurs initialement retenus, qu'un exercice pédagogique d'apprentissage par l'épreuve d'une démarche évaluative. Parmi beaucoup d'autres, les travaux de K. Thelen ont montré que les impacts exogènes (changement politique, réforme institutionnelle) produisent d'autant plus d'effets de changement qu'ils permettent l'actualisation et le renforcement de logiques déjà l'œuvre. C'est dans son modèle d'analyse la « sédimentation institutionnelle » qui permet de comprendre les possibilités et l'ampleur des changements observables. Dans le cas de l'évaluation, le choc exogène que constituerait la création d'une Haute autorité produirait d'autant d'effets qu'il aurait été précédé d'un travail d'apprentissage approfondi par la formation théorique et pratique.

2.5. L'évaluation comme expertise : la place du secteur privé

Les dynamiques observées précédemment (processus de décentralisation, l'enchevêtrement des compétences, le mouvement de rationalisation du secteur public, l'europeanisation des politiques publiques) ont conjointement

suscité une demande croissante d'évaluation au sein des administrations, centrales comme territoriales. Cette demande a contribué au développement d'un nouveau marché de l'expertise en évaluation. D'abord réalisé par des hauts fonctionnaires issus généralement des grands corps et amenés progressivement à solliciter le soutien ponctuel d'experts issus des milieux académiques (sciences de l'organisation en particulier), l'expertise en évaluation a progressivement été confiée à des cabinets de consultants privés. Ce glissement a imposé petit à petit la théorie que l'évaluation « n'est pas seulement une idée, mais aussi une industrie » (Angelmar, 1984). Si la demande des administrations et des collectivités a pu nourrir et structurer l'offre émanant du secteur privé, nous verrons comment à son tour cette offre est parvenue à susciter une demande et établir une forme de dépendance des acteurs publics aux opérateurs privés, à travers les différents chantiers de réforme de l'Etat (Matyjasik, Méasson, 2006).

Même si les données empiriques manquent sur le poids et l'organisation de ce secteur privé, les recherches récentes ont privilégié deux directions d'étude distinctes mais congruentes et toujours actives à ce jour. Une première approche, inspirée des apports de la sociologie du travail, s'est intéressée à la structuration et aux compétences de ces milieux du conseil et de l'expertise. Ces recherches ont analysé les compétences nécessaires au métier d'évaluateur, participant ainsi à l'institutionnalisation de l'évaluation dans le champ professionnel de l'expertise. Une seconde approche, en lien avec les travaux sur la réforme de l'Etat s'est attachée à analyser le rôle des experts et des cabinets de conseil dans les processus de modernisation des administrations. L'accent a été mis à la fois sur le recours croissant aux entités privées dans les dispositifs successifs de réforme et sur les implications de ce lien.

Professionnalisation et organisation d'un marché de l'expertise

Face aux besoins des élus et de l'administration, on assiste à partir des années 80 dans la plupart des pays occidentaux à la stabilisation d'un marché du conseil et d'expertise. Au sein de ce marché, dominé par les approches de stratégie organisationnelle et de gestion, l'évaluation ne représente qu'un segment minoritaire mais constitue, pour les cabinets de conseil qui s'y déploient, un levier intéressant pour pénétrer les administrations publiques, et diversifier les portefeuilles de clients.

Les travaux de Denis Saint-Martin, en popularisant la notion de « consultocratie » (Saint Martin, 1999), ont mis en évidence la montée en puissance des firmes de conseil sur la scène publique. Cet auteur montre comment suite à l'émergence du « nouveau managerialisme » dans les années 80, les gouvernants ont accru le recours aux services des consultants dans la réforme de l'Etat. Il explique en particulier que « la demande de conseil des administrations centrales ou locales [...] s'est développée à la faveur d'un retour en force de l'idéologie libérale dans les années quatre-vingts, mais aussi sous la pression d'exigences d'efficacité dont l'origine est plus large. Le thème de la modernisation de l'État a permis la pénétration du conseil dans l'administration de la même façon que le thème de la modernisation de l'économie française et des entreprises avait favorisé l'expansion du conseil dans les entreprises dans les années cinquante et soixante » (Saint-Martin, 2006).

Le rôle des grands cabinets de conseil et d'audit s'accroît également sous l'effet de la concentration rapide que connaît, en interne, le secteur dans la première moitié des années 90. Cette concentration conduit à l'émergence de multinationales du conseil qui développent des filiales spécifiques « services publics » et des prestations évolutives. Ces entreprises, d'abord centrées sur l'informatique, puis sur les audits comptables et les réformes de la comptabilité publique en viennent enfin à proposer leurs services pour des opérations de « réorganisations / débureaucratisations » (Berrebi-Hoffmann, 2002).

L'entrée de ces acteurs privés ne tient pas seulement aux stratégies, aussi offensives soient-elles, des opérateurs privés. Elle repose également fortement sur des formes de sollicitations émanant de la sphère publique elle-même. Comme le souligne N. Matyjazick, l'action publique a défini en grande partie l'étendue et la nature du mandat assigné aux cabinets de conseil. A travers les processus successifs de décentralisation, les transferts de pouvoir vers le local, l'enchevêtrement des niveaux de compétence, le poids grandissant des collectivités locales ou mouvement de rationalisation du secteur public ont contribué, de façon cumulative, à réunir les conditions favorables pour l'émergence d'une demande d'évaluation au sein de l'administration. En d'autres termes, ce marché doit en grande partie sa formation et le périmètre qu'il acquiert progressivement au double mécanisme d'ouverture politique du système, d'une part, et

d'institutionnalisation de l'évaluation au sein de l'appareil d'Etat, d'autre part.

Toujours est-il, qu'alors qu'elle était à peu près nulle il y a une vingtaine d'années, la part du secteur public est estimée aujourd'hui à près de 20 % du marché du consulting en Europe, évalué autour de 60 milliards d'euros (Saint-Martin, 1999). Bien que croissante, la part de l'administration dans le marché du consulting en France reste plus modeste que dans les autres pays occidentaux, autour de 8 %, comparée à 30 % pour le Royaume-Uni. L'industrie des consultants en gestion est largement dominée par les entreprises américaines qui représentent plus de la moitié du marché mondial pour les services de consultants.

Cela étant, l'univers du conseil est loin d'être homogène. Avec des caractéristiques socio-économiques, des activités et des profils de consultants différents, les sociétés de consulting forment un ensemble fragmenté au sein d'un espace mercantile composite. Pour reprendre l'expression utilisée par Odile Henry, ces sociétés sont plutôt comparables à des « supermarchés » (Henry, 1997).

La thèse de Nicolas Matyjazick dresse l'inventaire des cabinets privés intervenant dans le secteur public en France (encart n° 16). Cette liste met en évidence la taille relativement réduite de ces structures : près des trois quart d'entre elles compte moins de 20 salariés. Elle pointe également le caractère hétérogène des structures et la grande disparité entre les toutes petites entreprises (moins de 5 consultants) et les services évaluation plus étoffés hébergés par les grands groupes internationaux. Une autre caractéristique concerne la jeunesse de ces cabinets, la plupart ayant été créés dans les années 90, c'est-à-dire au moment où les réflexions sur la réforme de l'Etat et les transformations des administrations ont commencé à s'intensifier. Enfin, ces cabinets sont concentrés sur deux pôles géographiques : Paris et Lyon. Si la proximité des administrations centrales et des cercles d'influence explique l'implantation des cabinets dans la capitale, l'ancrage à Lyon est justifié le plus souvent par la possibilité d'intervenir dans le sud de la France tout en bénéficiant de l'exposition et de la visibilité offerte par la 2^{ème} ville de France.

Si l'on s'attarde sur le cas français, la genèse de ce type d'expertise permet de mieux en comprendre la morphologie actuelle. L'expertise privée en évaluation puise ses racines dans les milieux académiques et quelques

services ministériels pionniers (culture, équipement, éducation...). Le cas du laboratoire CEOPS (Conception d'Evaluation pour les Organisations et les Politiques publiques), créé à Lyon en 1988 par le Ministère de l'équipement en fournit un des exemples les plus connus (Barbier, Matyjazick, 2010)⁴¹. Ce laboratoire de l'Ecole Nationale des Travaux Publics de l'Etat (ENTPE) qui réunit une dizaine de chercheurs (essentiellement des politistes et des sociologues) se tourne à partir de la fin des années 90 vers les activités de conseil après avoir réalisé pendant plusieurs années un corpus de travaux d'accompagnement méthodologique pour les administrations et les institutions européennes. D'autres centres de recherches spécialisés dans le champ des politiques publiques (tels que le CERAT – Centre de recherche sur le politique, l'administration et le territoire - à Grenoble) de la sociologie urbaine (par exemple le LARES) ou de l'aménagement du territoire vont également s'intéresser à l'évaluation en tant que prestataires de service. Cet intérêt ne se manifesterait toutefois pas jusqu'à donner naissance à des cabinets de conseil ou des bureaux d'études spécifiquement dédiés aux activités d'évaluation. Tout en assumant leur « multipositionnalité », les chercheurs concernés veilleront en effet à rester en dehors du jeu en préservant leurs positions d'« experts » à la fois indépendants et sectoriels (Massardier, 1996).

Par ailleurs, comme le souligne N. Matyjazick, au sein des cabinets privés, l'évaluation des politiques publiques ne constitue que rarement une activité unique voire principale. L'évaluation représente en effet une intervention « secondaire » des cabinets de conseil derrière le conseil en management ou en stratégie et le conseil en développement territorial. Seulement une minorité de cabinets (7 entités en 2010) ont l'évaluation comme dominante exclusive au travers de travaux d'évaluation proprement dit, d'assistance à maîtrise d'ouvrage ou de formation à l'évaluation. Au contraire, la majorité des cabinets privés mènent en parallèle de leurs interventions évaluatives des activités de conseil et d'ingénierie auprès du secteur public, dans des champs sectoriels diversifiés (développement économique, développement territorial, politiques d'emploi et d'insertion ou encore politiques sociales), ou apparaissent comme des spécialistes du conseil en management.

⁴¹ Cf. *supra* 1.3 p.34

Encart n° 16 : Les cabinets de conseil en évaluation en France

En 2006, près de 150 cabinets de conseil intervenant sur la scène évaluative étaient recensés, dont une cinquantaine en Ile-de-France et une vingtaine en Rhône-Alpes. L'évaluation, qui ne représente que rarement la principale activité, est menée parallèlement aux activités de conseil (diagnostic, études....) dans le champ du développement économique, du développement territorial, des politiques d'emploi et d'insertion... Les cabinets sont souvent de faible taille (moins de 20 salariés) alors que les grands cabinets anglo-saxons (big five), dont le poids est limité, se concentrent surtout sur l'évaluation des fonds européens, plus rentables et rentrant dans des standards internationaux qu'ils maîtrisent davantage. Notons l'opposition entre ceux qui font primer les considérations méthodologiques et ceux qui privilégient au contraire la connaissance de la politique évaluée. L'évaluation territoriale a progressivement émergé, mais en abritant des conceptions et pratiques disparates.

(Source : N. Matyjazick, 2010)

La recherche d'une forme de reconnaissance professionnelle et la formalisation des compétences spécifiques au métier d'évaluateur constituent un second traceur pour appréhender l'organisation du secteur privé. En Europe, ce type d'approches, relativement peu développées, peut se lire en filigrane dans la définition des standards d'évaluation définis au niveau européen à partir des années 90 (cf. 2.3).

La création d'un « ordre » des professionnels de l'évaluation est également une problématique récurrente au sein de la communauté des évaluateurs. L'exemple du Canada est régulièrement cité. La Société canadienne d'évaluation (SCE), constituée en tant qu'ordre professionnel pour l'évaluation rend ainsi possible l'accréditation des personnes réunissant les compétences scolaires et professionnelles nécessaires pour assurer des services et prestations d'évaluation (encart n°17). La création de ce titre d'évaluateur accrédité (ÉA) a été conçu en vue de « définir, reconnaître et faire la promotion d'un exercice professionnel qui répond aux plus hauts standards d'éthique, de qualité et de compétence par le biais d'un programme de titres professionnels » (SCE, 2010).

Encart n° 17 : Le titre d'évaluateur accrédité (ÉA) au Canada

Le titre est un service offert par la SCÉ à ses membres. Ceux-ci peuvent décider, sur une base volontaire, de devenir évaluateurs accrédités. Le titre reconnaît les personnes qui ont la scolarité et l'expérience nécessaire pour fournir des services d'évaluation et instaure une culture de l'apprentissage continu au sein de notre communauté en mettant en œuvre des exigences de maintien et de renouvellement des qualifications.

Tous les évaluateurs agréés doivent être membres en règle de leur ordre professionnel et sont régis par un code de déontologie très strict qui dicte leur démarche professionnelle et qui les oblige, par exemple, à produire un rapport complet et détaillé qui accompagne et motive chacune de leurs conclusions de valeur.

De plus, chaque évaluateur doit détenir des assurances responsabilités professionnelles pour ainsi offrir une sécurité essentielle à toute sa clientèle.

Enfin, quiconque ne peut s'improviser évaluateur agréé, car ce dernier doit détenir une formation adéquate qui débute d'abord par un baccalauréat en administration des affaires avec une concentration en immobilier. Vient ensuite une année complète de stage auprès d'une firme d'évaluation accréditée et finalement l'examen de l'Ordre des Évaluateurs Agréés du Québec.

(Source : Site de l'Ordre des Évaluateurs Agréés du Québec ; <http://www.oedq.qc.ca/>)

En France, les responsables des différents cercles d'évaluation, publics comme privés, se sont pour le moment toujours refusés à appliquer de tels principes. Les réflexions ont davantage porté sur l'élaboration de standards communs et de référentiel de compétences spécifiques au métier d'évaluateur. À travers l'édiction d'une Charte (adoptée en 2003 et révisée en 2006) ou le recensement de normes de qualité, la SFE a ainsi développé une série de réflexions sur les standards de qualité et de déontologie des systèmes évaluatifs. Ces réflexions, conduites en collaboration avec l'Association Finances-Gestion-Évaluation des Collectivités Territoriales (AFIGESE-CT) et l'organisme France Qualité Publique prenant le relai des travaux menés par le CNE jusqu'à la fin des années 1990. Elles s'inscrivent également dans une logique de consolidation et d'autonomisation de l'identité de l'évaluation par rapport aux démarches voisines (audits de performance, management par la qualité...).

La thèse de Claire Tourmen soutenue en 2007 constitue également une source d'informations intéressante pour aider à circonscrire les compétences spécifiques aux évaluateurs. Elle distingue plus précisément trois types de

métiers liés à l'évaluation : les responsables d'évaluation en administration publique, les chargés d'évaluation externes et les responsables d'évaluation qui sont en même temps chargés d'évaluation ou d'auto-évaluation au sein de leur propre administration. Ces métiers ont en commun qu'ils appellent la mobilisation de neuf compétences génériques : des compétences aussi bien cognitives, opérationnelles que sociales (encart n°18). La déclinaison et le croisement de ces compétences doit permettre de constituer un référentiel des métiers et des compétences qui fait aujourd'hui défaut dans la communauté des évaluateurs.

Encart n° 18 : Les neuf compétences génériques des évaluateurs

A. Cognitives :

- Connaissance des mécanismes de décision et de mise en œuvre des politiques publiques dans différentes administrations
- Connaissance des secteurs à évaluer et de leurs logiques (possible spécialisation sectorielle)
- Connaissance des systèmes d'acteurs, à savoir des mécanismes institutionnels entre les acteurs et des compétences respectives des institutions ((Etat, Région, Commission européenne...))
- Connaissance des principes méthodologiques et déontologiques de l'évaluation de politiques publiques (historique, paradigmes, principales méthodes de recherche en sciences sociales...)

B. Opérationnelles :

- Savoir exploiter des bases de données : utilisation des bases de données internes aux administrations, analyse et capitalisation des données existantes
- Savoir utiliser des outils d'animation du travail collectif (techniques d'animation de réunion, outils de gestion de projet...)
- Savoir utiliser des techniques d'analyse de données couramment utilisées et/ou propres à l'évaluation (techniques d'entretien, de questionnaire, d'analyse statistique, d'analyse multicritères...)

C. Sociales :

- Maîtrise de la relation commerciale : démarcher la clientèle, connaissance de la concurrence, impératif de performance
- Maîtrise de la relation aux élus et aux services opérationnels : gérer les attentes, connaissance des relations de pouvoir, impératif d'opérationnalité

(Source : Claire Tourmen, 2007)

La place et le rôle des consultants dans les réformes de l'administration

Dans une veine différente, des travaux de recherche se sont intéressés aux acteurs privés du conseil et de l'évaluation pour analyser comment au cours des vingt dernières années, les consultants en gestion sont devenus les acteurs de plus en plus visibles des processus de réforme du secteur public (Denis Saint Marin, 2006).

De manière générale, le recours aux cabinets privés va connaître une accélération à partir du début des années 80. Sous l'influence des idées du « nouveau management public » (NMP) et de la montée de la « nouvelle droite », les gouvernements à la recherche d'une plus grande efficacité ont commencé à dépendre des consultants extérieurs pour le transfert des idées et des pratiques de gestion du monde des affaires vers le secteur public. Au Royaume-Uni, alors que dans l'année suivant l'élection de Margaret Thatcher, le gouvernement dépensait environ six millions de livres en services de consultants, ce chiffre s'élevait à 246 millions à la fin de son mandat de premier ministre en 1990. Au Canada, sous le gouvernement conservateur, les dépenses de consultants sont passées de 56 millions de dollars en 1984 à presque 190 millions de dollars en 1993 (Saint Martin, 1999).

Si en France, l'existence d'institutions comme l'École nationale d'administration (ÉNA) et des grands corps a généralement eu pour effet de restreindre l'accès des consultants au marché de l'administration centrale des exemples témoignent, le recours à l'expertise privée a, semble-t-il toujours été perçue plutôt positivement par les membres de l'exécutif.

Historiquement, l'arrivée au pouvoir d'une coalition de droite en 1986 va donner aux consultants une visibilité plus grande au sein de l'État. En 1986, deux consultants du secteur privé sont nommés au cœur de l'appareil d'État pour conseiller le gouvernement Chirac sur la réforme de l'État. Tous deux issus d'Eurequip, G. Raveleau et H. Sérieyx sont nommés, le premier au cabinet du ministre de l'économie, et le second auprès du ministre de la fonction publique. Les deux consultants se voient confier la responsabilité de développer une politique de promotion de la qualité et de l'innovation. Un rapport en ce sens est remis au gouvernement sur la base duquel celui-ci adopte en 1987 sa politique de « qualité et d'innovation ». Cette politique

traduit une volonté d'adaptation à la culture administrative du savoir-faire du privé.

Mais c'est surtout à partir des années 2000 et la succession dans un temps court des dispositifs de rationalisation de l'administration que va s'imposer le recours crescendo aux instruments du *New Public Management* et aux opérateurs du secteur privé et, ce faisant la diffusion des standards de la nouvelle gestion publique. Trois phases peuvent ainsi être distinguées (Bezes, 2010).

1 - La période d'élaboration et de mise en œuvre de la LOLF (1998-2006) est marquée tout d'abord par l'importation tangible et systématique, de part et d'autre (Parlement et Budget), de techniques néo-managériales diffusées par l'OCDE, cette fois adoptées dans un cadre opérationnel : celui de la réforme de la procédure et des formats budgétaires et des processus de gestion des finances publiques. Les administrateurs civils du Sénat, qui négocient la réforme, mobilisent largement les fascicules de l'OCDE. Mais ce corpus co-existe avec les expériences françaises antérieures défendues par la Direction du budget : la mise en place de démarches objectifs-moyens de contractualisation au ministère de l'équipement (depuis les années 80) et au sein de la direction générale des impôts (à la fin des années 90) sont précisément analysées. De son côté, l'inspection générale des Finances remet, en mars 2000, une étude de benchmark sur les systèmes de gestion publique par la performance et les procédures budgétaires dans huit pays (Canada, Danemark, Grande-Bretagne, Pays-Bas, États-Unis, thématiques et en deviennent les porteurs. Dans cette première phase, le recours aux cabinets de conseil est écarté mais l'argument qui le justifie est significatif. Comme l'indique Jean Picq⁴² alors en charge d'une mission de modernisation de la fonction publique : « *Je ne voulais pas de cabinet d'audit. C'était une réflexion sur l'État, pas une réforme administrative. En revanche, je les ai associé deux ou trois fois comme consultants sur ma méthode et ma matrice. On voulait puiser dans les forces de l'État. On a choisi par cooptation ou par appel au*

⁴² Conseiller maître à la Cour des Comptes, Jean Picq est chargé en 1994 par le Premier ministre Edouard Balladur d'une mission de réflexion sur les « responsabilités et l'organisation de l'Etat » en France. Il s'agit d'une véritable charte de la réforme de l'administration française où il est question du statut des fonctionnaires et de leurs fonctions électives, de la gestion du personnel administratif, etc.

corps » (Bézes, 2002). L'expertise reste alors dominée par les hauts fonctionnaires des grands corps.

2 - Dans une deuxième phase, qui se produit entre 2005 et 2007, la création de la DGME et le lancement des audits de modernisation de l'Etat font massivement appel aux cabinets de conseil en stratégie en matière de sources et de contenus d'expertise. La formule qui sera systématisée et étendue dans la RGPP repose sur la réalisation d'audits par des membres des inspections des administrations associés à des consultants, avec la volonté de « croiser les expertises » et de s'appuyer sur des acteurs extérieurs. Portée par le ministère du budget, cette stratégie de réforme de l'État se fonde sur deux partis pris. Le premier est l'idée que les solutions de réforme doivent venir de l'extérieur et ne pas être laissées aux seuls exercices d'introspection internes à l'État. Le deuxième parti pris concerne le contenu des audits. La commande n'est plus d'abord celle de réflexions d'ensemble sur les dysfonctionnements et la réorganisation de l'État débouchant sur des instruments ou des réformes à introduire. Elle devient une attention systématique accordée à la dépense sur le modèle explicite de la revue des programmes pratiquée au Canada en 1995 et articule systématiquement des propositions de réformes de structures à la réduction des dépenses publiques ou l'accroissement des performances. Les différentes commissions qui sont installées à l'époque (2004-2006) dans le cadre du Commissariat à la réforme de l'Etat (Commission Camdessus, Commission Pébereau) mêlent de multiples groupes d'acteurs (lobbies patronaux, instituts de recherche sur les administrations publiques, *think tanks* et acteurs du privé) dont plusieurs d'entre eux occuperont des postes éminents dans les cabinets ministériels quand il s'agira d'appliquer cette feuille de route. Ces commissions jouent un rôle important de traduction et de compromis : leurs rapports rendent compatibles les principes d'organisation du système administratif à la française avec les préceptes du *New Public Management (NPM)*, largement présents dans les réformes étrangères et dans les milieux d'expertise transnationaux, qu'ils combinent sous le label de « l'État-stratège ». Les instances centrales de réforme de l'État (délégation interministérielle à la réforme de l'État, puis DGME) ont utilisé les entreprises de conseil dont le rôle s'est également amplifié dans le cadre de la LOLF. L'implication des cabinets à culture anglo-

saxonne (Accenture ; PriceWaterhouse-Coopers ; Cap Gemini ; Bearing Point regroupant KPMG consulting et Arthur Andersen France) ou même des structures françaises (IDRH) est alors croissante.

3 - La mise en œuvre de la RGPP qui compose la troisième phase s'inscrit pleinement dans cette perspective qu'elle systématise en confiant aux cabinets de conseil, particulièrement aux consultants en stratégie, un rôle inédit dans le contexte français. Pour reprendre le cadre analytique de Denis Saint-Martin, les élites politiques et les groupes restreints de hauts fonctionnaires qui les entourent sont à la fois convaincus de la nécessité de réformes de restructuration et prêts à s'appuyer fortement sur des consultants (Saint-Martin, 2001). De manière inédite en France, les équipes constituées pour les audits RGPP (environ 200 personnes) associent officiellement les membres des inspections générales et des consultants externes dont l'expertise repose sur une légitimité différente de celle des hauts fonctionnaires. L'implication des inspections est un premier fait remarquable. Pour l'Inspection générale des finances, il a été calculé que la RGPP avait absorbé 34 % du temps des inspecteurs du service entre juillet et décembre 2007, douze inspecteurs généraux ayant été nommés chefs d'équipes sur un total de 21 missionnés par le Premier ministre, assistés de 35 inspecteurs des finances (IGF, 2008). Par contraste avec les audits de modernisation, dont une grande partie est mise en ligne, les audits RGPP sont restés des documents confidentiels. Peu impliquée en amont dans le contenu des audits RGPP, les activités de la DGME prennent un tour nouveau dans la phase de mise en œuvre qui fait de l'organisation un véritable lieu d'interface entre administrations et cabinets de conseil. A cette époque, la composition de la DGME est révélatrice de la place nouvelle des consultants. En 2009, sur 143 personnes travaillant au sein de ces trois départements, environ 46 % sont des contractuels recrutés par le biais de contrats à durée déterminée (deux ou trois ans), très souvent issus de structures de conseil privées, les autres membres étant fonctionnaires. Les rôles assurés par la DGME correspondent à ses fonctions d'interface. D'un côté, la DGME assure les fonctions de « greffier » du processus RGPP en assurant le suivi des décisions des différents conseils de modernisation des politiques et, surtout, de leur état d'avancement (voir en ligne la mise en scène des réalisations sous la forme de feux de

couleurs différentes). D'un autre côté, notamment par le biais de son service « Conseil » (présenté comme le « cabinet de conseil interne de l'administration ») et du département « Leviers » en son sein, elle joue un rôle d'appui aux projets de réorganisation des ministères affectés par la RGPP et capitalise sur l'expertise du secteur privé du conseil dont elle reprend en compte les démarches et le langage. Elle peut alors intervenir seule ou mener des missions mixtes, avec des consultants et sous la responsabilité d'un de ses membres. Elle peut aussi conseiller les ministères dans leur recherche d'un cabinet. Le service « Projet » est également en lien avec des consultants, notamment sur les systèmes d'information et l'administration électronique. Sur la base des rapports annuels de performance (RAP) concernant le programme 221 (Stratégie des de conseil des services « Innovation », « Conseil » et « Projets ») avoisinent les 45 millions d'euros. L'homogénéité des services de la DGME et des méthodes qu'elle diffuse ne doit pas être exagérée mais les interlocuteurs rencontrés soulignent qu'elle connaît un fort taux de rotation de son personnel et s'appuie sur des personnels jeunes.

Significativement, et sans qu'il soit possible de le chiffrer, les grands cabinets de conseil en stratégie jouent un rôle plus important qu'auparavant : McKinsey, BCG (Boston Consulting Group), à un moindre titre Roland Berger Strategy Consultants sont des cabinets privilégiés dans les missions de réorganisation. D'autres cabinets majeurs (Accenture, Cap Gemini, Ernst and Young), aux activités multiples (stratégie, audits et surtout systèmes d'information, branche particulièrement rémunératrice) sont également très impliqués dans les réorganisations. Le changement d'échelle est réel.

Comme l'indique P. Bezes, il est difficile de mesurer avec précision si les consultants solidifient, ou au contraire ébranlent, la capacité d'élaboration des politiques (Bezes, 2010). Cette implication des consultants dans les politiques de réforme de l'État a évidemment de puissants effets sur le contenu et les formes des politiques de réforme. Dans le cas de la RGPP, les consultants « amènent des méthodologies et des savoir-faire relatifs au suivi du changement (restitutions (*reporting*), tableau de bord, remontée d'information, bilan d'avancement) mais aussi aux formats de présentation (le diaporama de type *powerpoint* prenant la place du rapport, plus caractéristique de la culture administrative française). Leur influence ne doit

ni être exagérée ni minorée. Sur le fond, il apparaît que les membres des inspections dans les équipes d'audit ont souvent cadré et limité leurs rôles. On peut faire l'hypothèse que les relations ont été variables d'une mission à l'autre ». (p.15) Toutefois, la présence systématique des cabinets d'audit favorise évidemment la diffusion de savoirs de réorganisation spécifique.

Pour ce qui est des effets sur le gouvernement, les opinions semblent assez contradictoires. Pour D. Saint-Martin (2006), « d'un côté, les fonctionnaires et les consultants prétendent que le recours aux consultants dans le processus administratif est une manière d'affermir la capacité d'élaboration des politiques des gouvernements en mobilisant une expertise souvent inexistante à l'interne. De plus, les consultants sont considérés comme porteurs de souplesse : on les fait intervenir à la discrétion du ministère ou de l'organisme concerné, et on peut les choisir pour répondre à certains besoins spécifiques. Mais de l'autre côté, des critiques suggèrent que les consultants affaiblissent les capacités du gouvernement en usurpant peut-être des fonctions politiques. Ainsi, les consultants, s'apparentant en cela à d'autres conseillers externes, appartiendraient à un 'gouvernement fantôme' qui aurait effectivement enlevé aux élus et aux bureaucrates la fonction d'élaboration des politiques » (p. 755).

Finalement, depuis 2007, l'État a fortement accru le recours aux consultants privés dans ses réorganisations internes et ses recherches d'économies. Le bilan de la RGPP réalisé dans un rapport parlementaire confié à deux députés en décembre 2012 soulignait ainsi que Bercy avait consacré plus de 102 millions d'euros entre 2006 et 2011 à ces prestations, qui ont bénéficié à tous les ministères dans le cadre de la RGPP. Dans une partie consacrée à la direction générale de la modernisation de l'État (DGME), le service chargé à Bercy d'accompagner les ministères dans leurs réorganisations, les parlementaires révèlent le détail des montants versés par le ministère du Budget à des consultants privés entre 2006 et 2011 (encart n° 19).

Encart n° 19 : Le bilan des prestations confiées aux cabinets privés dans la RGPP

Dépenses « consultants » pour la modernisation de l'Etat (en M€)

Consultants (Prog. 221)	2006 (exéc.)	2007 (exéc.)	2008 (exéc.)	2009 (exéc.)	2010 (exéc.)	2011(exéc.)	2012 (prévision exéc.)	2013 (PLF)
<i>Administration électronique</i>	22,74	29,17						
<i>Modernisation gestion</i>	9,96	11,49						
<i>Pilotage transversal/supports</i>	3,95	4,03						
<i>Qualité et simplification</i>	1,68	3,33						
Conseil			29,16	18,43	21,23	19,40	16,35	16,04
Innovation			0,53	2,72	1,92	2,50	2,19	2,25
Projets			31,63	21,28	22,51	21,55	14,43	14,09
Actions transverses			3,83	2,49	2,40	3,59	4,53	4,05
TOTAL	38,33	48,02	65,15	44,92	48,06	47,04	37,50	36,43
Evolution n-1/n		25,28%	35,67%	-31,05%	6,99%	-2,12%	-20,28%	-2,85%

Source : Secrétariat général du MINEFI et DGME.

Conclusion : bilan de l'institutionnalisation

Des scènes « ouvertes »

Au terme de ce panorama des démarches d'évaluation déployées en France et dans les démocraties occidentales, le constat peut-être fait à la fois d'une grande continuité et d'un formidable éparpillement dans les tentatives successives d'institutionnalisation et de concrétisation de l'évaluation.

Les instances et les controverses qui se nouent encore aujourd'hui autour de l'évaluation montrent que cette dernière a non seulement été intégrée par les différents acteurs mais qu'elle constitue en outre un creuset où s'inventent et s'expérimentent, dans la confrontation directe aux enjeux de terrain, de nouvelles pratiques et représentations.

L'analyse historique des différentes configurations de développement de l'évaluation atteste qu'il y a bien eu des scènes multiples où l'évaluation a pu se déployer jusqu'à se concrétiser dans les pratiques et acquérir un certain degré de pérennisation. Des débats initiaux menés sur les finalités de l'évaluation aux montages institutionnels qui ont essayé de les rendre opérationnels en passant par les instruments juridiques tendant à normer les pratiques, voire à les rendre obligatoires, l'ensemble de ces mobilisations a contribué à diffuser ce type d'activité et à assurer sa continuité. La pérennité, voire à certains moments, l'intensification des débats et controverses ont aussi participé à l'institutionnalisation et la légitimation de l'évaluation. Lorsque l'on analyse les transformations observables depuis les années 60, il n'est pas difficile de montrer l'approfondissement tant des démarches que des idées et des représentations, en particulier dans les milieux administratifs et politiques. Des avancées importantes ont eu lieu sur le plan des méthodes, en particulier dans la capacité à différencier l'évaluation des démarches voisines : audit, contrôle, inspection. Les formes d'évaluation ont été différenciées : évaluation des programmes ex ante/ ex post ; évaluation accompagnement-adaptation ; auto-évaluation et apprentissage. Enfin, de nouveaux acteurs ont été sollicités, l'évaluation-expertise (réservée aux spécialistes) est aujourd'hui complétée par des démarches collectives (ouvertes aux diverses parties-prenantes) voire participative (associant de bout en bout les parties-prenantes). Ces divers changements ont aussi permis

le développement d'un langage et de catégories de pensée commune, sous l'influence de la production de standards de bonnes pratiques. L'Union européenne a contribué à cette dynamique à travers l'évaluation des fonds structurels mais aussi par la diffusion de ressources méthodologiques et la mise en réseaux des acteurs nationaux de l'évaluation. L'évaluation a également été progressivement saisie par le niveau territorial (en particulier les Régions) qui y ont vu l'opportunité d'asseoir leurs compétences par rapport à l'Etat central. Ainsi a été effectué un rapprochement des points de vue sur l'évaluation. Mais le déploiement de l'évaluation s'est aussi traduit sur le plan opérationnel par la généralisation de cette pratique au sein des administrations et des collectivités, avec un recours de plus en plus important à l'expertise du secteur privé. Alors que jusque dans les années 1980, elle ne constituait encore qu'un objet intellectuel de curiosité, voire de débat, l'évaluation est aujourd'hui pratiquée dans la plupart des secteurs d'action publique, y compris ceux qui au départ y étaient les plus réticents : culture, santé, social...

En France, même si l'évaluation reste moins développée que dans les pays anglo-saxons, le bilan dressé par Jean-Claude Thoenig en 2005 résume bien la place et le positionnement de cette pratique :

« L'opinion des cercles avertis est souvent cyclothymique quand il s'agit de juger des avancées de la modernisation administrative. Tel est le cas pour l'évaluation en France. A l'euphorie sans limite succède dès le début des années 90 un scepticisme exagéré. Il est de bon ton d'en parler désormais comme d'un phénomène passé, d'en regretter le prétendu déclin et d'en accabler les illusions perdues [...] [mais] l'évaluation a fait son trou, même si certains espéraient mieux. Elle a réglé son compte aux ayatollahs de tous poils qui espèrent une gouvernance par la vérité des chiffres ou qui rêvent d'une agora évaluatrice. Elle a pris sa place dans le rang des techniques et démarches qui composent la palette du management public, plus précisément pour ce qui concerne la connaissance des conditions de management des politiques publiques. L'échec du rêve de 1990 aura masqué un bilan qui, tout compte fait, à ce stade, n'est pas déshonorant ». (J.-C. Thoenig, 2005, p. 122)

A un niveau plus macro, la stabilisation de l'évaluation peut également être montrée à partir des efforts de classification des activités. La typologie élaborée par S. Jacob et F. Varone en 2004 constitue une référence. En s'appuyant sur les travaux de J.-E. Furubo, R. Rist et R. Sandahl (2002), ces auteurs proposent une série d'indicateurs synthétiques qui croisent le degré d'institutionnalisation de l'évaluation et le degré de maturité des pratiques en ce domaine. Leur modélisation aboutit à la distinction entre trois grandes catégories de pays. D'un côté, ceux qui comme les Etats-Unis, l'Australie, le Royaume-Uni et, dans une certaine mesure la France obtiennent un indice élevé attestant d'une institutionnalisation stabilisée et d'une maturité de la pratique évaluative élevée. D'un autre côté, des pays comme la Belgique, la Nouvelle-Zélande, l'Espagne ou le Japon dans lesquels la maturité et le degré d'institutionnalisation sont faibles. Entre ces deux extrêmes figurent un certain nombre de pays de catégorie intermédiaire dans lesquels la pratique est répandue mais le degré d'institutionnalisation modéré. Le modèle Jacob-Varone n'identifie pas de pays ayant une faible maturité et forte institutionnalisation.

Cette approche met l'accent sur l'institutionnalisation envisagée comme un processus dynamique par lequel des dispositifs sont créés, modifiés, voire supprimés mais concourent, sur une période plus ou moins longue, à l'expansion et à la pérennisation de la pratique évaluative. Les opportunités de changement contribuent au final, et par adaptations incrémentales, à produire du sens autour de cette activité et à légitimer certains principes d'organisation et de travail plutôt que d'autres.

Pour autant, à y regarder de plus près, dans le cas Français, aucune des scènes classiques mentionnées précédemment (Cour des Comptes, Commissariat au plan, Parlement, Inspections administratives) n'est devenue le principal porteur de l'évaluation. Si chacune a participé à la mise en visibilité et à la concrétisation de l'évaluation, aucune d'entre elles n'a permis d'imposer un modèle de référence pour l'évaluation. Chacune reste aujourd'hui une scène ouverte où se produisent toujours, certes sous des formes qui ont évolué, des débats et discussions autour des finalités, des enjeux et des méthodes d'évaluation.

Six lignes de débat

Six lignes de débat classiques mais récurrents sont donc observables.

1 – L'évaluation est régulièrement invoquée comme un moyen de légitimation des programmes d'action publique, mais aussi comme un contre-pouvoir par rapport à l'exécutif. Certes, le contexte a changé : il ne s'agit plus de justifier les mesures de lutte contre la pauvreté (« *war on poverty* ») ni de mettre en œuvre les grands programmes d'armement. Mais dans une société en tension marquée par le contexte de crise économique durable (SFE, 2012), elle apparaît pour les décideurs comme une ressource précieuse pour démontrer l'utilité et l'efficacité des actions, tant en termes budgétaires que d'effets sociaux. C'est ce que rappelle M. Bovens lorsqu'il souligne que l'évaluation demeure dans la période contemporaine un enjeu pour un ensemble d'institutions qui revendiquent chacune des méthodes et des capacités d'analyse spécifiques pour pouvoir agir plus fortement sur le « *policy making* » (Bovens et al., 2006). En France, ce type de débat a été porté dès la fin des années 80 par le collectif de haut-fonctionnaires « Service public » qui voyaient dans l'évaluation et ses démonstrations d'utilité sociale, un contre-feu à l'encontre des pressions néo-libérales de réduction des budgets publics. Au niveau territorial, l'évaluation a accompagné les différentes vagues de décentralisation. De nombreuses collectivités ont vu dans l'évaluation un moyen de légitimer leurs politiques et se démarquer du pouvoir central. Aux différents échelons territoriaux, les assemblées se sont emparées de l'évaluation pour faire valoir leur légitimité à prendre des décisions et à asseoir leurs compétences spécifiques.

2 – L'évaluation reste un outil pertinent pour la compréhension des dynamiques d'action publique et des processus administratifs. Aux premiers travaux qui se sont tout particulièrement attachés à rendre compte des écarts entre l'idéal-type weberien de la bureaucratie et l'observation empirique des appropriations dispersées des programmes par les structures administratives, ont progressivement succédé des approches davantage centrées sur l'analyse des processus de transformation de l'Etat. L'évaluation a permis de ne plus penser l'administration en termes seulement institutionnels et juridiques, mais comme un acteur hétérogène inséré dans des réseaux et en interaction constante tant avec les acteurs économiques et sociaux qu'avec les différents niveaux politiques (territoriaux, nationaux et européens).

3 – les approches démocratiques, portées en France dès la fin des années 1980 avec le Rapport Viveret ont été diffusées, sous des formes adaptées, dans d'autres champs. Ainsi, les récentes réflexions sur la place de l'évaluation dans le travail parlementaire (MEC, CEP, LOLF...) montrent que l'impulsion démocratique initiale n'a jamais été réellement abandonnée. Même si la production de diagnostics et de jugements techniques et économiques sont les plus fréquents, l'évaluation reste toujours associée, plus ou moins explicitement, à sa capacité de validation démocratique des choix politiques. Plusieurs responsables politiques se positionnent ainsi en faveur d'un « changement de métier des députés » qui devraient « passer plus de temps au contrôle de l'action gouvernementale et moins au vote de la loi » (Copé, 2008). Sur le plan des démarches concrètes, la prise en compte croissante des « citoyens » dans les travaux d'évaluation montre que les approches démocratiques ont diffusé y compris au sein des approches rationalisantes de la modernisation de l'action publique. Après la phase des évaluations managériales, faites pour et par la seule administration, un certain nombre d'études et travaux réalisés depuis la fin des années 90 se sont efforcés d'associer des tiers. Cette nouvelle vague d'évaluation se caractérise par la mise en œuvre de démarches plus ambitieuses en termes d'implication de la société civile qu'il s'agisse des évaluations participatives et depuis quelques années des expérimentations d'évaluations « émancipatrices » (*empowerment evaluations*).

4 – L'importance d'une culture « sectorielle » de l'évaluation dans les administrations et les ministères a considérablement freiné la concrétisation du dispositif démocratique et de débat politique proposé par le rapport Viveret. On retrouve là une conséquence du phénomène général de prééminence du pouvoir exécutif inhérent à la Cinquième République. En France, l'essentiel des activités concrètes d'évaluation est, encore à ce jour, réalisé au niveau institutionnel des ministères. A fortiori depuis la suppression du Commissariat général du Plan, ils sont les principaux lieux d'expérimentation et de mise en œuvre concrète de ce type de production de connaissance. Ces évaluations faites dans les ministères présentent de nombreuses limites, en particulier dans le choix des programmes étudiés (rarement les plus problématiques, plutôt ceux qui valorisent l'action menée), la faiblesse de leurs moyens et leur très faible impact sur les choix budgétaires. Leur cloisonnement sectoriel, leur confinement au sein des arcanes administratives ainsi que le manque

d'échanges méthodologiques entre secteurs sont un frein au développement élargi et transversal de l'évaluation.

5 – L'évaluation technologique (*technological assessment*), malgré son caractère limité, demeure une des plus stables sur le plan institutionnel. Parmi l'ensemble des tentatives, les évaluations technologiques ont constitué une réponse à un ensemble de controverses récentes (gestion des déchets nucléaires, expositions aux faibles doses, nanotechnologies...). Les crises environnementales et d'aménagement ont été des facteurs favorisant ce domaine. De plus, sous l'influence du modèle scandinave des Conférences de consensus, la consultation, voire la participation du public ont été progressivement intégrées dans certaines démarches. Ainsi, en France comme dans d'autres pays (Allemagne, Belgique, Espagne), l'évaluation des choix technologiques est devenue au fil du temps une composante importante du dialogue social (Boy, 2000).

6 – Enfin, la performance publique a trouvé un nouveau souffle dans le cadre des audits RGPP aujourd'hui prolongés par la MAP annoncée en décembre 2012 par le Premier ministre Jean-Marc Ayrault. Cette continuité officialise le consensus politique autour de l'impératif de renforcement de l'efficacité des politiques et des services publics. Le Comité Interministériel de la Modernisation de l'Action Publique (CIMAP) chargé de sa mise en application fait de l'évaluation l'un de ses 5 axes d'intervention aux côtés de la simplification des démarches administratives et des normes (axe 1), la mesure de la qualité du service public (baromètre indépendant, axe 2), l'accélération de la transition numérique de l'administration (axe 3) et l'intégration des agences et des opérateurs dans la modernisation de l'action publique (axe 5). Le dispositif prévoit que l'ensemble des politiques publiques soit évalué avant 2017 et tous les ministères verront au moins l'une de leurs politiques publiques faire l'objet de travaux d'évaluation au cours de l'année 2013. Dans ses principes, la démarche d'évaluation proposée présente plusieurs avancées dans la culture administrative française de l'Etat : elle est propre à chaque sujet et non uniforme ; elle touche aux objectifs des politiques et non à la seule gestion des ressources et elle cherche à associer les parties prenantes plutôt qu'à miser sur la seule expertise des agents publics. Comme pour la RGPP, l'expertise des cabinets privés est sollicitée pour accompagner la démarche. Le besoin d'expertise s'avère cependant différent dans la mesure où il porte moins sur des questions d'organisation et de stratégie que sur la réalisation de diagnostics

et la remontée d'informations qualitatives sur la mise en œuvre des politiques. L'ouverture vers des démarches de prospective basés sur la construction de scénarios est également recherchée. Finalement, les approches en termes de performance publique apparaissent dominantes dans le paysage politique et la liaison entre évaluation et recherche d'efficacité / d'efficience est présentée ces dernières années comme le facteur le plus favorable au développement d'une culture de l'évaluation.

Si elle explique la continuité, la poursuite de ces lignes de débats et controverses témoignent aussi de la fragmentation voire de la fragilité de l'évaluation. Deux approches permettent de rendre compte de ce paradoxe. La première approche prend acte du caractère intrinsèquement indépassable de cet ensemble composite. Elle souligne la différenciation des modèles et des pratiques d'évaluation en fonction de plusieurs critères tels que leurs finalités (évaluation managériale / démocratique / participative...), leurs lieux de production (évaluation parlementaire / experte / indépendante...) ou encore les processus historiques ayant permis la diffusion de modèles ou de dispositifs spécifiques d'évaluation (dispositif technocratique / corporatiste / participatif / libéral...).

La deuxième approche insiste quant à elle sur l'incapacité des acteurs à trancher dans une direction ou bien dans l'autre, à se positionner en faveur d'une problématisation de l'évaluation. C'est cette thèse que nous pourrions développer sur le cas français pour conclure notre réflexion.

« Tous compétents ! » : L'indéfinissable approche à la Française

Dans leur approche institutionnelle de l'évaluation, F. Varone et S. Jacob accordent une place plutôt « flatteuse » à la France par rapport au degré d'institutionnalisation de l'évaluation. Mais ce résultat est trompeur car il est le reflet de la période pendant laquelle la recherche a été effectuée, le début des années 2000, période durant laquelle l'existence du Conseil scientifique de l'évaluation au sein du Commissariat général de plan dominait encore le paysage institutionnel. Le déclin rapide qui aboutira le 27 octobre 2005 par le remplacement du Commissariat au Plan par un Centre d'analyse stratégique était difficilement prévisible.

En France, comme dans d'autres pays, les difficultés d'institutionnalisation ont principalement été du à des problèmes du côté de « l'offre » (Jacob, 2005, p. 861). Pour des raisons intellectuelles autant que pour des logiques

d'appartenance institutionnelle, il n'y a jamais eu en France de véritable constitution d'une « communauté des évaluateurs » et l'approche « à la française » est demeurée floue et inclassable. Même si le modèle promu par le conseil scientifique de l'évaluation se différencie des systèmes d'audit et d'expertise close au profit de démarches collectives, ce modèle a rencontré deux types de limites. D'un côté, il est largement resté sans prise sur les évaluations menées au sein des ministères et dans les inspections administratives. D'un autre côté, les instances interministérielles mises en place par le CSE sont loin d'avoir adopté des démarches comparables qui auraient concrétisé un modèle de référence (Lascoumes, Setbon, 1996).

L'étude des phases d'avancée et de repli qu'a connu l'évaluation en France montre qu'il ne saurait y avoir d'institutionnalisation acquise. L'expérience de ces trois décennies d'institutionnalisation permet de mettre en évidence la caractéristique principale du système à la française : le fait qu'à la fois tous les acteurs concernés se considèrent comme compétents pour évaluer et qu'aucun n'est parvenu à imposer son approche ou sa pratique.

Les clivages entre hauts fonctionnaires, universitaires et responsables de bureaux d'étude sont demeurés très forts. L'opportunité d'ancrer la pratique de l'évaluation dans les pratiques administratives et politiques n'a jamais été vraiment saisie en raison des divisions et des querelles sur les démarches et les finalités de l'évaluation. Presque 30 ans après, il n'y a pas de consensus sur le but poursuivi : doit-il être à dominante managériale ou d'apprentissage démocratique (Duran, Monnier, 1992 ; Leca, 1992) ? Le débat est toujours ouvert. Cette situation a freiné le développement des pratiques, la stabilisation des choix méthodologiques et la constitution d'un « corps » d'évaluateurs.

Sur le plan institutionnel, il n'y a jamais eu d'arbitrages politiques sur l'existence d'une ou de deux structures centralisées. La recherche infructueuse d'un modèle d'évaluation au Parlement, depuis les propositions du rapport Viveret (1989) jusqu'aux créations récentes d'instances (Offices, MEC, CEP, CAP...) pas plus que l'absence d'accord sur la création d'une instance unique et centralisée (INSEE ? Commissariat général au Plan ? Centre d'analyse stratégique ? Autre AAI ?...) n'ont permis de parvenir à un accord sur une organisation unifiée et réellement opératoire.

L'évaluation est également restée un enjeu important de lutte pour les inspections. Sur la période 1970-1990, on observe des luttes et des coalitions

d'acteurs (association Services Publics, Club Cambon et aujourd'hui divers « think tanks » de droite comme de gauche) se réclamant de « l'école française » de l'évaluation. Les grands corps de l'Etat ont défendu leurs prérogatives en matière d'évaluation, chacun rappelant à l'envi qu'il était le plus compétent ou le plus outillé pour mettre en œuvre des évaluations. Le déroulement de la deuxième Conférence nationale de l'évaluation organisée en décembre 1993 conserve la trace de ces débats nourris entre le Conseil d'Etat et la Cour des Comptes (CES, 1993). A l'époque, Guy Braibant, alors Président de la section du rapport et des études du Conseil d'Etat indiquait que le Conseil d'Etat, « avait toujours fait de l'évaluation sans le savoir ou sans le dire ». Il prenait comme exemple les nombreux arrêts en contentieux effectués par les sections du Conseil d'Etat. Le représentant de la Cour des Comptes signalait quant à lui que les magistrats, de par leur indépendance, leur savoir-faire et leurs moyens techniques et humains (les 500 magistrats répartis dans les Cours des comptes régionales) disposaient d'une « force de frappe » inégalée pour mener à bien des évaluations. Il reconnaissait cependant à l'époque que la Cour s'était toujours interdite d'intervenir de manière autonome en matière d'évaluation et que l'entrée en contact avec les bénéficiaires finaux restait un obstacle majeur pour les instances gouvernementales ou législatives. Depuis cette conférence, les occasions se sont répétées où chacune des deux instances n'a jamais concédé à l'autre la possibilité (ou le droit !) de représenter l'organe officiel d'évaluation des politiques publiques.

Récemment, l'introduction de la réforme constitutionnelle de 2008 confiant à la Cour des comptes une mission d'évaluation des politiques publiques a fait l'objet de vives réactions de la part de députés qui souhaitaient plutôt voire confier ces missions de contrôle et d'évaluation au Parlement (encart n°20). A l'initiative d'un groupe de députés, un amendement a été déposé pour proposer la création d'un office parlementaire d'évaluation des politiques publiques. L'objectif était de contrer la montée en charge des activités d'évaluation de la Cour et de faire de l'Assemblée Nationale l'organe de référence en matière d'évaluation. Cette initiative aboutira cependant à la création une nouvelle fois d'instances en charge de l'évaluation.

Encart n° 20 : Tentative (manquée) de putsch parlementaire contre la rue Cambon (juillet 2008)

Un épisode intéressant s'est déroulé en juillet 2008 à l'occasion des débats parlementaires portant sur le projet de loi constitutionnelle n° 993 et concernant « la modernisation des institutions de la Ve République ». Malgré diverses réticences de la majorité, Louis Giscard d'Estaing, député du Puy-de-Dôme, vice-président de la commission des Finances, dépose à l'Assemblée nationale, le 8 juillet 2008, un amendement visant à introduire dans la Constitution le principe de doter le Parlement d'un office parlementaire d'évaluation et de contrôle des politiques publiques. N'ayant pas abouti, il le présentera à nouveau en deuxième lecture. Il n'aboutira pas. C'est à l'inverse un renforcement des missions de la Cour des comptes qui a été introduit dans la réforme constitutionnelle de juillet 2008 (article 47).

Cet amendement visait à compléter de la façon suivante l'article Article 21 du projet de loi : « Art. 47-3. - Le Parlement est doté d'un office parlementaire d'évaluation et de contrôle des politiques publiques. Une loi fixe les modalités de son fonctionnement. »

Ce projet de création d'un office parlementaire d'évaluation des politiques publiques suscite des débats intéressants, mais qui au bout du compte montrent l'impuissance du Parlement à se doter d'une structure d'évaluation des politiques publiques. Didier Migaud, président de la commission des Finances de l'Assemblée, opposé à cette initiative déclare « nous n'avons pas de culture du contrôle et de l'évaluation ; nous n'avons qu'une culture de la soumission et de la démission permanentes ».

Ce sont des députés, incarnant l'aile libérale de l'UMP, qui ont voulu rendre au Parlement la prérogative croissante de la Cour des comptes en matière d'évaluation des politiques publiques. Cette idée avait déjà été exprimée dans les recommandations du comité Balladur et reprise dans l'avant-projet de loi. La commission des Lois avait aussitôt corrigé le dispositif initial, en stipulant que le Parlement "vote la loi, contrôle l'action du gouvernement et concourt à l'évaluation des politiques publiques".

Cela n'était pas suffisant pour une cinquantaine de députés UMP, qui souhaitaient l'inscription dans la Constitution d'un office parlementaire de contrôle des politiques publiques. Ils ont annoncé leur intention de défendre un amendement en ce sens. "Nous voulons que le Parlement ait la capacité autonome de concourir à l'élaboration des politiques publiques, il s'agit d'un sujet majeur dans l'équilibre de nos institutions", plaide Louis Giscard d'Estaing. Selon Jean-Michel Fourgous, "comme il n'y a pas de contrôle efficient, la France est l'un des pays où l'on a la plus grosse progression de dépense publique". Ils se réfèrent aussi à la déclaration du futur Président de la République lors de son discours du 14 janvier 2007 « La démocratie irréprochable » qui énonçait : « ce n'est pas une démocratie où l'exécutif est tout et le Parlement rien. C'est une démocratie où le Parlement contrôle l'exécutif et a les moyens de le faire. » Selon cette logique, le Parlement devrait se doter d'un office parlementaire d'évaluation et de contrôle des politiques publiques. Ce point de vue n'est pas partagé par le Parti

socialiste, mais le député UMP des Yvelines souligne toutefois qu'il y a "consensus" avec le PS sur la nécessité d'un contrôle parlementaire... Cet amendement, également appuyé par le think tank libéral IFRAP, vise à atténuer la montée en puissance de la Cour des comptes qui inquiète tout autant Bercy que les parlementaires spécialistes des questions budgétaires.

L'organisme envisagé aurait une double structure. D'une part, un comité d'audit parlementaire (CAP), qui pourra prendre la forme d'une délégation, devant lequel aboutiront les rapports des divers organismes de contrôle afin de leur donner l'impact nécessaire. D'autre part, dans la mesure où, à la différence du modèle britannique, la Cour des Comptes ne peut être rattachée au Parlement, un organisme d'enquête assurera la liaison entre le CAP et les organismes de contrôle. Cette cellule regroupera des magistrats détachés de la Cour des Comptes, des Inspecteurs généraux de différentes administrations, et des conseils issus du secteur privé. Cette double structure n'induirait ainsi pas de coût supplémentaire. D'une part le CAP serait constitué d'une vingtaine de parlementaires. Quant à l'organisme d'enquête responsable des rapports d'autre part, il serait dirigé par deux parlementaires désignés par le CAP et secondés par des Inspecteurs généraux de l'administration ou des magistrats de la Cour des Comptes détachés. Le recours à des cabinets privés serait enfin réalisé dans le cadre des crédits existants des Assemblées, ainsi globalisés et coordonnés dans leur emploi.

A travers ces deux exemples, on mesure les différents facteurs qui ont conduit en France à la non-émergence d'un acteur fort de l'évaluation, que certains appellent pourtant de leurs vœux.

Enfin, sur un troisième niveau, on peut faire l'hypothèse que l'impératif délibératif et participatif qui s'est étendu depuis quelques années au champ de l'évaluation peut également, dans une certaine mesure, être lu comme une manifestation additionnelle de la dilution du leadership en ce domaine. Faute d'une légitimité suffisante, les acteurs se sont tournés vers les citoyens dits « ordinaires » pour fonder la légitimité de leurs activités d'évaluation. « Tous compétents même le profane ! » : ces tentatives d'association des tiers à la pratique évaluative a achevé d'ajouter un niveau de compétence supplémentaire sans que celle-ci soit connectée à un organe indépendant et représentatif capable d'en porter politiquement les résultats.

Cette conjonction de luttes de pouvoir et d'absence d'arbitrages explique finalement en grande partie qu'à côté des effets d'apprentissages observés en

termes d'acculturation et d'acclimatation à l'évaluation, il n'y ait pas eu toutefois l'émergence d'un modèle « français » (comme cela a été le cas dans d'autres pays) et que la pratique de l'évaluation soit restée superficielle.

Controverses et résistances autour du mythe de l'évaluation

Parallèlement à ces difficultés d'institutionnalisation, l'évaluation doit faire face, en toile de fond, à des critiques récurrentes qui remettent périodiquement en cause sa portée sinon sa légitimité. Plus que l'évaluation, ces critiques prennent pour cible le culte du résultat et la mobilisation à outrance des indicateurs de performance. Schématiquement, elles sont développées dans trois champs distincts. Nous pouvons en résumer les principaux argumentaires afin de saisir la tonalité et la portée des controverses et des résistances les plus couramment opposées à l'évaluation.

1 – Dans le champ intellectuel, tout d'abord, les débats philosophiques et métaphysiques contre la généralisation des procédures d'évaluation continuent d'occuper une place importante. Le trait commun de ces critiques anti-évaluation porte sur la définition du périmètre des activités d'évaluation et les conditions de leur acceptabilité sociale. Ce qui est en cause généralement c'est le fait de mesurer et de classer la performance des individus. L'évaluation y est entendue au sens large : sondage d'opinion, notation, techniques économétriques, statistique, classements, certification, indices bibliométriques... Les critiques les plus vives sont formulées dans les domaines où l'évaluation est considérée comme la moins naturelle, en particulier l'éducation, le travail, l'enseignement ou encore la recherche. Les réactions qu'a soulevé en France le fonctionnement de l'Agence de l'évaluation de la recherche et de l'enseignement supérieur (AERES) sont indicatifs des arguments les plus couramment formulés contre l'évaluation. Mise en place en 2007, l'AERES a fait l'objet de vives critiques quant à ses méthodes et sa tendance à assimiler évaluation et notation. Le bureaucratisme de l'instance tout comme le coût exorbitant de son fonctionnement ont également nourri les réquisitoires des observateurs favorables à sa suppression. Un rapport remis en janvier 2013 au Premier ministre par le député socialiste Jean-Yves Le Déaut dresse le bilan du

fonctionnement de l'AERES⁴³. Si le principe de l'évaluation n'est pas remis en cause, le rapport préconise en revanche le remplacement du système de notation par un dispositif d'évaluation basé sur la production d'avis motivés. En appelant à une simplification des procédures d'évaluation, le rapport conclut également que « trop d'évaluation nuit à l'évaluation » (p.96). Dans un autre domaine, les enquêtes Pisa menées par l'Organisation de coopération et de développement économique (OCDE) ont suscité de nombreuses réactions de la part des spécialistes des sciences de l'éducation. Ces enquêtes, qui comparent régulièrement les performances des élèves de 15 ans dans une trentaine de pays ont été fortement critiquées du fait de leur logique du classement extrêmement réductrice ainsi que de l'absence de prise en compte du contexte culturel dans l'appréciation des compétences scolaires. Prenant appui sur ce type d'exemples, une abondante littérature s'est attachée à décrire les effets pervers et les abus liés à la « banalisation des procédures d'évaluation » (Zarka, 2009). Pour les psychanalystes, philosophes et sociologues qui la dénoncent, l'essentiel des accusations se construit autour d'un l'argumentaire tendant à dénoncer le culte de la performance et l'évaluation comme cheval de Troie du *new public management* et, plus largement de la doctrine néolibérale (encart n° 21).

Encart n° 21 : L'idéologie et la tyrannie de l'évaluation

Depuis quelques années, le développement de l'évaluation a été accompagné par des nombreuses publications (Zarka, 2009 ; Del rey, 2013 ; Gori, 2011 ; Vidaillet, 2013...) dans le champ de la philosophie, de la psychologie ou des sciences de l'éducation et du travail prenant pour cible la mesure de la performance et ses avatars : classements, notations, baromètres, certifications.... Schématiquement l'argumentaire déployé contre l'évaluation (entendue au sens large) peut être résumé selon les idées suivantes :

- il existe une idéologie de l'évaluation, à savoir une norme puissante qui s'impose à l'ensemble de la société (Zarka, 2009) ;*
- cette idéologie est une imposture, dans la mesure où elle résulte d'un processus de construction et d'un travail historique de reconnaissance de la position d'évaluateur (Gori, 2011) ;*

⁴³ J.-Y. Le Déaut, (2013) « Refonder l'Université, dynamiser la recherche : mieux coopérer pour réussir », Propositions de transcriptions législatives et réglementaires des conclusions des Assises 2012 de l'enseignement supérieur et de la recherche, Rapport à Jean-Marc Ayrault, Premier ministre, le 14 janvier 2013.

- *les gouvernants abusent de cette évaluation et elle renforce le pouvoir jusqu'à devenir « tyrannique » : l'évaluation, diffuse et continue, profite de la transparence et les évaluateurs sont cachés (Del Rey, 2013) ;*
- *les outils techniques sont la principale arme des évaluations : ils sont porteurs de cette tendance à la simplification du réel qui permet de donner aux résultats des évaluations une apparence d'objectivité (Zarka, 2009) ;*
- *en définitive, cette « fièvre » de l'évaluation (Damon, 2009) signe l'émergence d'une société malade, qui bégaie, incapable de penser autrement que par fragments ;*

2 – Une seconde forme de résistance concerne l'absence de reconnaissance scientifique de l'évaluation. Le positionnement de l'évaluation reste en effet problématique dans les milieux académiques. Les trois disciplines que sont l'économie, la sociologie et la science politique entretiennent des rapports complexes avec l'évaluation.

Tandis qu'elle a bénéficié à ses débuts du parrainage de plusieurs courants de recherche en sciences sociales (sociologie des organisations, analyse des politiques publiques, sciences administratives...) l'évaluation est depuis marginalisée par les chercheurs en science politique et en sociologie qui s'en sont progressivement détournés. Les arguments invoqués pour expliquer cette rupture insistent sur les décalages entre le temps de la recherche et le temps de l'évaluation. Mais l'absence de passerelles entre les activités de recherche et les activités d'expertise ou de consultance et sa faible reconnaissance, pour ne pas dire son ignorance, entre pairs expliquent également cette mise à distance de l'évaluation. L'inadéquation des centres de recherche à répondre aux commandes des collectivités et des administrations dans certaines expériences a sans doute contribué pour sa part à éloigner plusieurs d'entre eux au profit des acteurs du secteur privé (cabinets, bureaux d'études...). Toutefois, plusieurs chercheurs continuent d'appeler de leurs vœux à un intérêt pour l'évaluation, par nécessité empirique comme par souci de mieux prendre en compte l'implication des citoyens dans l'élaboration et la conduite des politiques publiques (Faure, Smith, 2004).

Le positionnement des sciences économiques démontre un rapport différent entre cette discipline et l'évaluation mais qui reste à ce jour circonscrit et limité. On observe tout d'abord entre l'évaluation et l'économie une

continuité de pratiques qui vont du calcul économique public aux travaux plus récents de chercheurs se réclamant de l'économie publique (Barbier, Matyjazick, 2010). Historiquement, les relations paraissent avoir été nouée en ligne directe même si le label « évaluation des politiques publiques est minoré par les économistes qui privilégient la référence à l'aide à la décision publique » (Greffé, 1997). Plus qu'une branche de la science économique, l'évaluation apparaît comme un cadre d'analyse pour la modélisation des questions d'évaluation. Les travaux d'économie publique s'étant intéressés à l'évaluation portent ainsi une attention particulière à l'éclaircissement des techniques d'évaluation que sont notamment le taux de rendement socio-économique et les formules d'évaluation multi-critères (Baslé et Malgrange, 2006). Depuis une période récente, le rapprochement entre économie et évaluation se matérialise en particulier dans le retour des techniques d'évaluation expérimentale aléatoire nées dans les années 1960 dans le champ de la psychologie sociale aux Etats-Unis.

Si donc la sociologie et la science politique furent présentes dans les premiers travaux d'évaluation et les sciences économiques absentes, la tendance semble s'être inversée. La seconde est aujourd'hui devenue dominante et les techniques économétriques justement sollicitées. Dans les trois cas, il importe de noter l'absence de dialogue inter-disciplinaire alors que l'évaluation en tant que démarche d'investigation ouverte et transversale pourrait être un espace de rencontre de ces différentes disciplines.

3 – Sur le plan politique et institutionnel enfin, le potentiel critique de l'évaluation constitue un frein à son développement. Si en principe s'est construit au fil du temps un consensus sur les bénéfices de l'évaluation de l'action publique (au nom notamment d'une plus grande rationalité et d'une meilleure efficacité de l'intervention publique), dans les faits de nombreuses politiques continuent d'échapper aux évaluations. Cela tient en particulier au fait que « *l'enjeu de l'évaluation n'est pas seulement d'apprécier l'efficacité mais aussi d'imputer des responsabilités* » (Leca, Thoenig, 1997). Dans de nombreux cas, l'évaluation reste perçue par les décideurs comme porteuse d'un risque politique. Dans l'analyse qu'ils font des cas de non-évaluation, D. Gaxie et P. Laborier repèrent plusieurs difficultés (Gaxie, Laborier, 2003). La première est d'ordre technique : certaines politiques publiques ont des objectifs tellement généraux, diffus, voire polymorphes qu'il est bien difficile d'en apprécier les résultats. Tel est l'exemple de la politique de démocratisation culturelle, de prévention de la délinquance ou des politiques

d'intégration des populations d'origine étrangère. Dans ces politiques, la définition d'indicateurs et la recherche de relations de causalité se prêtent mal à un exercice d'évaluation. La seconde raison concerne les obstacles sociaux. Comme le rappellent les deux auteurs, l'évaluation s'inscrit toujours dans un rapport de force exempt de neutralité et qui amène à privilégier les intérêts de groupes sociaux par rapport à d'autres. Il n'est pas rare ainsi que les objectifs de « préservation de la paix sociale » (c'est l'exemple des dispositifs d'emplois jeunes ou des politiques dites de « zones prioritaires ») priment sur l'identification des solutions optimales pour la collectivité prise dans son ensemble (p. 214). Troisièmement, les limites au recours à l'évaluation sont également de nature cognitive. Pour apprécier la réussite ou l'échec d'une politique publique l'évaluation est en concurrence avec d'autres instruments ou sources de connaissance. La réussite d'un programme peut par exemple tenir plus dans ses effets symboliques et médiatiques que dans la mesure de son impact effectif. Les effets puissants des réalisations tangibles que l'« on voit » sur le terrain (construction d'un hôpital, ouverture d'un Zenith...) ou des mesures dont « on parle » dans les médias peuvent avoir valeur d'évaluation. Le succès médiatique ou symbolique de certaines actions peut ainsi dans certains cas se substituer aux besoins d'évaluation. Pour les promoteurs des « Journées du Patrimoine » par exemple, « les retombées médiatiques de l'évènement valent évaluation complémentaire et attestent de la réussite de l'action engagée aux yeux des protagonistes » (p. 215). L'empilement de dispositifs publics montre aussi que les décideurs « font » sans se préoccuper de leur efficacité : l'action publique est perçue comme « fonctionnant » car les décideurs accumulent les réalisations dans le domaine. Outre ces difficultés, le contournement des évaluations peut également résulter de stratégies objectives de contournement ou d'évitement, ce que D. Gaxie et P. Laborier qualifient « d'intérêts à la non-évaluation » (p.212). Ces stratégies sont à rapprocher des théories d'évitement du blâme (« *blame avoidance* ») popularisées par Kent Weaver (1986). Ces théories s'appuient sur l'observation que les individus (en l'espèce les décideurs) accordent plus d'importance aux pertes qui les affectent qu'aux gains qu'ils peuvent percevoir. Appliquées à l'évaluation, elles mettent en évidence le fait que les décideurs cherchent à éviter ou à contourner les évaluations afin de ne pas avoir à justifier leurs choix ou leurs raisons d'agir. Evaluer consiste la plupart du temps à introduire une incertitude qui peut-être vécue comme menaçante. Cette incertitude l'emporte le plus souvent sur l'attente d'un satisfecit. Ces

exemples amènent enfin à insister sur l'importance de la dynamique processuelle que comporte toute politique publique. Cette dernière montre qu'au-delà des résultats attendus des actions, le souci d'affirmer ou d'afficher une volonté de mettre fin à une situation indésirable l'emporte généralement sur la détermination à agir selon un diagnostic précisément établi. La dimension symbolique de certaines mesures (par exemple « la sécurité », la santé...) conduit ainsi les décideurs à privilégier les politiques ayant le plus d'effets sur les perceptions que sur l'efficacité attendue au regard du problème ou de l'enjeu identifié. Cela explique aussi que les enquêtes de satisfaction soient le plus souvent privilégiées par rapport aux études permettant d'apprécier l'efficacité spécifique de la politique menée.

Dans ce contexte, il n'est donc pas rare de voir un rapport d'évaluation (quand il a été commandé) étouffé ou des procédures évaluatives redéfinies en cours de route. Deux exemples concrets peuvent être donnés pour illustrer ce propos. Le premier concerne la mise en sommeil d'un rapport de l'INSERM par un Ministre en place (encart n°22). Le second, relaté par X. Molenat (2011) porte sur l'expérience « tronquée » d'évaluation *ex-ante* du revenu de solidarité active (RSA). Avant d'être adopté, le dispositif RSA a été expérimenté dans une trentaine de départements. Le but était de savoir s'il fonctionnait, quels devaient être son montant et les dispositifs d'accompagnement à prévoir. Mais les critiques méthodologiques formulées à l'encontre de l'expérimentation (en particulier sur le choix des départements) ont été telles que l'expérience a été interrompue au bout d'un an. Loin de se traduire par l'approbation et la généralisation du dispositif en fin d'expérimentation comme il était prévu initialement, l'expérience a été interrompue au bout d'à peine un an... par l'adoption et la généralisation anticipées du RSA. L'évaluation aura donc été loin de livrer tous les enseignements qu'elle promettait et qui auraient permis, selon ses défenseurs, de mettre en place un dispositif optimal

Encart n° 22 : Psychothérapies : le rapport d'évaluation censuré

En février 2004, l'Institut national de la santé et de la recherche médicale (Inserm) publie une expertise collective sur les psychothérapies. Trois approches sont comparées : les thérapies « psycho-dynamiques » ayant la théorie psychanalytique comme référence, les thérapies cognitivo-comportementales et les thérapies « familiale et de couple ». Sur toute une série de troubles mentaux (schizophrénie, troubles anxieux, autisme, hyperactivité...), le rapport d'évaluation met globalement en évidence une efficacité supérieure des thérapies cognitivo-comportementales et, dans une moindre mesure, des thérapies familiales. Selon ses auteurs, l'efficacité des thérapies psycho-dynamiques, peu étudiée, est rarement prouvée.

Les réactions à cette évaluation sont très vives, en particulier dans les milieux psychanalytiques. Si certains refusent le principe même de l'évaluation en santé mentale, d'autres pointent des méthodes biaisées. Le champ de l'expertise est ainsi limité aux maladies (alors que la santé mentale dépasse largement ce cadre). Elle laisse par ailleurs totalement de côté les consultations en cabinet privé. Enfin, le fait de prendre comme indicateur l'amélioration de l'état du malade (par rapport à un autre malade n'ayant pas suivi de thérapie) empêche de voir que, bien souvent, la prise en charge thérapeutique du patient n'apporte pas d'amélioration mais empêche que son état n'empire.

En février 2005, invité au Forum des psys, le ministre de la Santé Philippe Douste-Blazy affirme devant une audience de professionnels de la santé mentale que « le premier devoir d'une société est de reconnaître qu'il n'existe pas une seule réponse à la souffrance psychique », souffrance qui n'est « ni évaluable ni mesurable ». Il annonce également à son auditoire qu'il a fait retirer le rapport de l'Inserm du site internet du ministère de la Santé et qu'ils « n'en entendront plus parler ». L'Inserm, les auteurs de l'expertise, ainsi que de nombreuses associations de professionnels et de patients en santé mentale protesteront en vain contre cette censure.

(Source : X. Molenat, 2011)

Ces exemples démontrent bien que, tout en apparaissant pour les gouvernants comme un outil d'aide à la décision synonyme de rationalité et de modernité de l'action publique, l'évaluation est aussi porteuse, du fait de son exposition et de ses incertitudes, d'effets pervers et de dangers qui, dans la pratique, freinent son développement auprès de bon nombre de décideurs. Bien qu'étant officiellement présentée comme un réflexe naturel, ce réflexe reste rarement mis en œuvre et non appliqué dans de nombreuses circonstances.

En conclusion, le développement de l'évaluation entretient des rapports ambigus par rapport à ces trois foyers de critiques. D'un côté, les moments successifs de reformulations et d'interpolations entre les différents acteurs décrits précédemment puisent dans ces controverses et reconfigurent le paysage institutionnel au gré des résistances. D'un autre côté, les entreprises d'institutionnalisation étudiées tout au long de cette note paraissent relativement déconnectées de ces critiques. Aucune de ces controverses n'explique en effet la diffusion superficielle de l'évaluation, qui tient davantage des mécanismes institutionnels et politiques propres à l'administration française. Chacune d'entre elles compose cependant un arrière plan qui imprègne toutes les références à son développement et à son institutionnalisation. Toutes les trois ont en commun de mettre l'accent sur les insuffisances, les biais et les effets pervers de l'évaluation. Ce faisant, elles continuent à assurer sa permanence à l'agenda politique. Tout porte dès lors à penser que l'évaluation fonctionne comme les mythes décrits par Georges Dumézil : tout contribue à l'entretenir y compris sa dénonciation.

* * *

Bibliographie

- Agulhon C. (2006), « La mesure en éducation, un outil au service du politique », *L'orientation scolaire et professionnelle*, 35, n°4, p. 315-366.
- Amiot M. (1986), *Contre l'Etat, les sociologues : éléments pour une histoire de la sociologie urbaine en France, 1900-1980*, Paris, Editions de l'EHESS.
- Angelmar R. (1984), « L'évaluation : un outil d'apprentissage ? », *Politique et Management Public*, vol. 2, n°2, p. 115.
- Auvolat M., (1996), « L'évaluation de la politique des Conseils Régionaux : quelques enseignements tirés de l'expérience Rhône-Alpes »
- Barbier J.C. (2010), « Eléments pour une sociologie de l'évaluation des politiques publiques en France », *Revue française des affaires sociales*, n°101, pp. 25-49.
- Barbier J.-C., Matyjazick N. (2010), « Evaluation des politiques publiques et quantification en France : des relations ambiguës et contradictoires entre disciplines », *Revue française de socio-économie*, n°5, pp. 123-140.
- Baron G., Monnier E. (2003), « Une approche pluraliste et participative : Coproduire l'évaluation avec la société civile », *Revue « Informations sociales*, n°110, sept-2003.
- Barthe Y., (2002), *Le rôle de l'OPECST, une mise en perspective historique, Rapport pour le programme « Concertation, décision, environnement »*, CERAT-CSI, Grenoble.
- Baslé M., Pelé F. (1994), « L'évaluation des politiques publiques et des contrats de plan Etat-Région en France : la situation fin 1994 », *Les cahiers économiques de Bretagne*, n°4.
- Baslé M., Malgrange P., (2006), « Présentation générale. Aide à la décision publique : développements récents », *Economie et Prévision*, n°175, pp. 1-6.
- Bernard Bruhnes Consultant (en partenariat avec l'Association des Administrateurs Territoriaux de France), « Vingt ans de décentralisation. Où en est aujourd'hui le management public territorial ? » *Enquête sur les pratiques managériales dans les collectivités territoriales*, Septembre 2002,

n°10, 64 p.

http://www.brunhes.com/publications/cahiers/cahier10/Cahier_10_GBB.pdf

Bezes P. (2002), « La “mission Picq” ou la tentation de l’architecte. Les hauts fonctionnaires dans la réforme de l’État » in Nay O., Smith A. (dir.), *Le gouvernement du compromis. Courtiers et généralistes dans l’action publique*, Paris, Economica, coll. Études Politiques, 2002, pp. 111-147.

Bezes P. (2004), « Rationalisation salariale dans l’administration française : un instrument discret » dans Lascoumes P., Le Galès P., *Gouverner par les instruments*, Paris, Presses de Sciences Po, 2004, pp. 71-122.

Bezes P. (2009), *Réinventer l’Etat. Les réformes de l’administration française (1962-2008)*, PUF, Paris.

Bezes P. (2010), « ‘Morphologie de la RGPP’. Une mise en perspective historique et comparative », *Revue française d’administration publique*, n°136, pp. 775-802.

Bovens M., Hart P.T., Kuipers S. (2006), “The politics of policy evaluation”, *Oxford Handbook of public policy*, Oxford, chap 15, 319-335.

Bourdin J., André P., Plancade J.-P. (2004), *Rapport d’information sur l’évaluation des politiques publiques en France*, Sénat, rapport n°392, 30 juin.

Boy D., Donnet Kamel D., Roqueplo Ph, (2000), *Revue française de science politique*, 50, 4-5, pp. 779-810

Brooks H., (1973), « Le processus d’évaluation de la technologie », *Revue internationale des sciences sociales*, 25, 3, pp. 271-282.

Carson R. (1962), *Silent Spring*, New York, Shackleton.

Chanut V. (2002), « L’évaluation : affaire d’Etat ou question d’organisation ? », *Politiques et management public*, n°20-4, pp.1-32

CEPP (1998), *Séminaire franco-britannique sur l’évaluation de politiques publiques*, CSE - Tavistock Institute, 29-30 janvier 1998.

CES (1993), *Deuxième conférence nationale de l’évaluation*, 9 décembre 1993, Paris.

- Chelimsky E. (1985), « L'évaluation des programmes aux États-Unis », *Politiques et Management Public*, n°2, p199-213.
- Claret de Fleurieu P. (1995), « L'expérience de l'OPECST » in Delcamp A. et al. edit., *Contrôle parlementaire et évaluation*, Documentation française, Notes et études, n°5012-13, pp.75-79.
- Club Cambon, (1990), « Compte rendu de la journée d'étude sur l'évaluation des actions publiques, principes opératoires ? », *Politique et Management Public*, 8, 1, pp. 75-147.
- Cohen S. (1982), « Prospective et politique étrangère : le centre d'analyse et de prévision du ministère des Relations extérieures », *Revue française de science politique*, n°6, pp. 1055-1076.
- Cornut-Gentile F., Eckert C. (2011), Assemblée Nationale, Comité d'évaluation et de contrôle, « Une évaluation de la RGPP : méthode, contenus, impacts financiers », Rapport d'information, n° 4019, enregistré le 1^{er} décembre 2011
- Crozier M. (1963), *Le phénomène bureaucratique*, Paris, Seuil.
- Damon J. (2011), « La fièvre de l'évaluation », *Sciences humaines*, 15/06/2011, (http://www.scienceshumaines.com/la-fievre-de-l-evaluation_fr_24277.html).
- Delcamp A., Bergel J.L., Dupas A. (1995), *Contrôle parlementaire et évaluation*, Documentation française, Notes et études, N°5012-13.
- Deleau M. (dir.) (1986), *Évaluer les politiques publiques : méthodes, déontologie, organisation, rapport du groupe de travail "Méthodes d'évaluation des politiques publiques"*, Commissariat général du Plan, Paris, La Documentation française.
- Delley J.-D. (2006) in Lascoumes P. et al, *L'évaluation des politiques publiques entre enjeu politique et enjeu de méthode*, CEVIPOF, Sciences Po, 123 p.
- Del rey A (2013), *La tyrannie de l'évaluation*, Paris, Broché.
- Derouet J.L. (2006), « Entre la récupération des savoirs critiques et la construction du management libéral », *Revue française de pédagogie*, n°154, p. 5-18.

Derthick M. (1972), *New towns in town, why a Federal Program failed*, Washington, The Urban Institute, 102p.

Desrosières A. (1993), *La politique des grands nombres. Histoire de la raison statistique*, Paris, La Découverte.

Donzelot J., Viveret P. (1989), « Evaluation et démocratie, contrôle des politiques publiques et formation de l'opinion », *Esprit*, octobre, p. 42-56.

Duran P., Monnier E. (1992), « Le développement de l'évaluation en France ; Nécessités techniques et exigences politiques », *Revue française de science politique*, pp. 235-262.

Dubois V. (1999), *La politique culturelle. Genèse d'une catégorie d'action publique*, Paris, Belin.

Duran P. (1993), « Les ambiguïtés politiques de l'évaluation », *Pouvoirs*, Paris: PUF, n° 67.

Duran P., Monnier E. (1992), « Le développement de l'évaluation en France ; Nécessités techniques et exigences politiques », *Revue française de science politique*, avril 1992, pp. 235-262.

Duran P. (2010), « L'évaluation des politiques publiques : une résistible obligation », Introduction générale, *Revue française des affaires sociales*, n°1-2, pp. 5-24.

Ellul J. (1964), *The technological society*, New York, Knopf.

Epstein R. (2010), « À quoi sert l'évaluation ? Les leçons de la politique de la ville », *Tracés. Revue de Sciences humaines*.

Epstein Renaud, (2007), « L'évaluation en développement ? Retour sur vingt ans d'évaluation de la politique de la ville » in B. Bouquet, I. Sainsaulieu et M. Jaeger éd., *Les défis de l'évaluation en action sociale et médicosociale*, Paris, Dunod, pp. 75-93.

Faucher-King Fl., Le Galès P. (2010) « Auditer la société », *Les gouvernements New Labor*, Sciences Po Les presses, pp. 89-94.

Faure A., Smith A. (2004), « L'évaluation objet de savoir et enjeu de pouvoir », *Sciences de la société*, n°32, pp. 149-152.

- Fetterman D., Kaftarian S.J., Wandersman A. (eds.), (1995), *Empowerment Evaluation : Knowledge and Tools for Self-Assessment and Accountability*, SAGE Publications.
- Fontaine J. (1986), « L'évaluation des politiques publiques par le Conseil Régional de Bretagne », Communication pour l'atelier : l'évaluation des politiques publiques par les Conseils régionaux. Discussions d'expériences en cours, Ve Congrès AFSP, Aix-en-Provence.
- Fontaine J. (1996), « Evaluation des politiques publiques et sciences sociales utiles. Raisons des décideurs, offres d'expertise et usages sociaux des pratiques régionales », *Politix*, n°36, pp. 51-71.
- Furubo J.-E., Rist R., Sandahl R. (eds), 2002, *International Atlas of Evaluation*, New Brunswick & Londres, Transaction Publishers.
- Gaxie D., Laborier P. (2003) « Des obstacles à l'évaluation des actions publiques et quelques pistes pour tenter de les surmonter », P. Favre et al. *Etre Gouverné*, Presses de Sciences-Po, chap. 10, pp. 201-221.
- Gazier B. (2000), *Tous « sublimes », vers un nouvel plein emploi*, Paris, Flammarion, 2003.
- Girard A. (1971), *Comptes culturels de la nation*, Paris, La documentation française, pp. 2-3.
- Gori R. (2011), *La Folie Evaluation : les nouvelles fabriques de la servitude*, Paris, Broché.
- Greffe X. (1997), *L'évaluation des projets publics*, Paris, Economica.
- Gremion P. (1976), *Le pouvoir périphérique, bureaucrates et notables dans le système politique français*, Paris, Seuil.
- Guba E., Lincoln Y. (1987) « The contentances of fourth generation evaluation: description, judgement and negotiation » in Palumbo D., (dir.), *The politics of program evaluation*, Los Angeles, Sage.
- Hargrove, E.C. (1975) *The Missing Link : the Study of the Implementation of Social Policy*, Washington, The Urban Institute.

Henry (O.), « La construction d'un monde à part. Processus de socialisation dans les grands cabinets de conseil », *Politix*, vol. 39, n°3, 1997, pp. 155-177.

Horaist J. (1992) (Dir.), *Partenaires pour évaluer. Organiser l'évaluation des politiques publiques locales*, Paris, Commissariat général du Plan, p.129-138.

Hitch Ch. J., Mc Kean R. N. (1960), *The economics of defense in the nuclear age*, Washington, Rand Corporation.

Ion J., Peroni P., Roux J. (1996), *Evaluation, évaluabilité et action publique*, Cresal, Rapport au Commissariat Général du Plan.

Jacob S., Varone F., Genard, J.-L. (Dir.) (2007), *L'évaluation des politiques au niveau régional*, Bruxelles, Peter Lang, Coll. Action publique.

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, n° 5-6, p. 835-864.

Jobert, (B.), Muller, (P.) (1987), *L'Etat en action, politiques publiques et corporatismes*, Paris, PUF.

Jones C.O. (1970), *An Introduction to the Study of Public Policy*, Belmont, Duxbury Press.

Kessler M.-C., Lascoumes P., Setbon M., Thoenig J.-C. (Dir.) (1998), *L'évaluation des politiques publiques*, Actes de colloque, 5-6 février 1997, Ecole normale supérieure, Cachan, Paris, L'Harmattan

Kirszbaum T., Epstein R., (2005), *Synthèse nationale des évaluations à mi-parcours des contrats de ville 2000-2006*, Saint-Denis - La Plaine, Rapport pour la DIV.

Knoepfel P., Larue C., Varone F. (2001), « L'évaluation des effets des politiques publiques », *Analyse et pilotage des politiques publiques*, Helbing, & Lichtenhahn, Genève, pp. 253-287.

Lacaze A. (2005), « La LOLF : simple outil de management, ou dogme écrasant ? », *Gérer et comprendre, Annales des mines*, n°81, pp. 5-13.

Laget J.-J., Benamour M. (1996), « Une expérience d'évaluation de politique locale : l'exemple du département de l'Hérault », *Politiques et management public*, vol. 14 n° 2, p. 105-127.

Lagroye J., Bastien F., Sawicki F. (2002), « Les formes d'institutionnalisation », *Sociologie politique*, Presses de Sciences-Po – Dalloz, 4^e edit. p. 528-534

Lassave P. (1997), *Les sociologues et la recherche urbaine dans la France contemporaine*, Toulouse, Presses universitaires du Mirail.

Lascoumes P., Setbon M. (1996), *L'évaluation pluraliste des politiques publiques: enjeux, pratiques, produits*, Paris, GAPPENRS/Commissariat général du plan.

Lascoumes P., Le Galès P. (2007), *Sociologie de l'action publique*, Paris, Armand Colin, collection 128, 126p.

Lascoumes P., Varone F., en collaboration avec Pons X. (2006), *L'évaluation des politiques publiques entre enjeu politique et enjeu de méthode*, CEVIPOF, Sciences Po, 123 p.

Laswell H. (1956), *The decision Process : Seven Categories of Functional Analysis*, College Park, University of Maryland Press.

Laufer R., Paradeise, C. (1982), *Le Prince bureaucrate*, Paris, Flammarion.

Leca J. (1993), « Sur le rôle de la connaissance dans la modernisation de l'Etat et le statut de l'évaluation », *Revue française d'administration publique*, n°66, pp. 185-196.

Leca J. (1993), « L'évaluation dans la modernisation de l'Etat », *Politiques et management public*, Volume 11, n°2, juin, pp. 161-173.

Leca J. (1992), « Développement ou crise d'identité de l'évaluation », Conseil scientifique de l'évaluation, *L'Evaluation en développement*, Documentation française, pp. 1-12.

Leca J., Thoenig J.-C. (1997), « L'évaluation, pour gouverner autrement », *Le Monde*, 09/12/1997.

Lobet-Maris C., Kusters B. (1992), « Technology Assesment : un concept et des pratiques en évolution », *Technologie de l'information et société*, 4, 4, pp. 435-455.

Mairate A. (2007) « La Commission européenne et l'évaluation régionale » in Jacob S., Varone F., Genard, J.-L., *L'évaluation des politiques au niveau régional*, Bruxelles, Peter Lang, Coll. Action publique.

Matyjazick N. (2010), « L'évaluation des politiques publiques dans une France décentralisée. Institutions, marché et professionnels », Thèse pour le doctorat de sciences politiques, Université Montesquieu –IEP Bordeaux.

Matyjazick N., Measson L. (2006), « L'évaluation des politiques publiques locales : opportunité professionnelle et effets sur l'action territoriale », CERVL IEP Bordeaux, Journées d'études organisées par l'Association Française de Science Politique, juin 2006

Maury E. (1998), « Le Parlement français face au défi de l'évaluation des politiques publiques », in Kessler et al., *L'évaluation des politiques publiques*, Paris, L'Harmattan.

Massardier (G.), « Les savants les plus demandés. Expertise, compétences et multipositionnalité. Le cas des géographes dans la politique d'aménagement du territoire », *Politix*, n°36, 1996, p. 163-180.

Molenat X. (2011), « La fièvre de l'évaluation », *Sciences humaines*, 15/06/2011, (http://www.scienceshumaines.com/la-fievre-de-l-evaluation_fr_24277.html)

Mayne J., Bemelmans-Videc M.L., et al (1992), *Advancing public policy evaluation, learning from international experiences*, Amsterdam, Elsevier.

Mironesco Ch. (1997), *Un enjeu démocratique : le Technology assessment*, Genève, Georg Edition.

Moatti J.P. (1983), « L'expérience américaine de l'évaluation technologique », *Culture technique*, 10, p. 211-221

Monnier E. (1987), *L'évaluation de l'action des pouvoirs publics*, Paris, Economica.

Nioche J-P., Poincard R. (1984), *L'évaluation des politiques publiques*, Paris, Economica.

Normand R., Derouet J.L., (2011), *Le développement d'une culture de l'évaluation dans l'Education nationales : comment les enseignants utilisent-ils les résultats des évaluations nationales ?*, INRP, Rapport pour le Commissariat général du plan, 195 p.

Perret B. (1999), « L'évaluation des politiques publiques dans les administrations d'Etat, éléments de diagnostic », *Congrès de la SFE*, Marseille, 4-5 juin.

Perret B. (2001), *L'évaluation des politiques publiques*, Paris, La découverte.

Pfieger S., (2010), « Augustin Girard, les études au service de la décision politique », *Hermès*, 56, p. 192-20.

Petrella R., (1992), « Le printemps du Technological Assessment en Europe, faits et questionnements », *Technologies de l'information et société*, 4, 4 p. 425-434.

Saint-Martin D. (1999), « Les consultants et la réforme managérialiste de l'Etat en France et en Grande Bretagne : vers l'émergence d'une « consultacratie » ? », *Revue canadienne de science politique*, vol. 32, n°1, mars 1999, pp. 41-4

Saint-Martin D. (2001), *Building the New Managerial State. Consultants and the Politics of Public Sector Reform in Comparative Perspective*, Oxford University Press.

Saint-Martin D. (2006), « Le consulting et l'Etat : une analyse comparée de l'offre et de la demande », *Revue française d'administration publique*, n° 120, pp. 743-756.

Société Canadienne d'Evaluation (2010), Programme de titres professionnels : guide du postulant, mai.

Spentheuer V. (1995), *L'évaluation de politique, usages sociaux, 3 études de cas d'évaluation*, L'Harmattan, 1995.

Spentheuer (1998), « L'évaluation des politiques publiques en France, avatar de la planification », Thèse pour le doctorat de sciences politique, Université Pierre Mendès-France, Grenoble II.

Spenehauer V., Warin P., (2000), « L'évaluation au service des conseils régionaux », *Sociologie du travail*, vol. 42, no 2, p. 245-262.

SFE, Mouterde F., Trosa S. (2010), « Les nouvelles frontières de l'évaluation - 1989-2009 : vingt ans d'évaluation des politiques publiques en France, et demain ?, Paris, L'Harmattan

SFE, Baron G., Matyjazick N., (2012), « L'évaluation des politiques publiques : défi d'une société en tension », Paris, L'Harmattan.

Sfez L. (1970), *L'administration prospective*, Paris, A. Colin

Taiclet A.-F. (2001), « Legitimacy and Accountability in Multi-Level Games. An Empirical Assessment through Local Economic Development Policies in France », Communication, 29th Joint Sessions of Workshops, Grenoble, ECPR, 2001.

Tenzer N. (dir.) (1995), *L'évaluation : une nécessité stratégique pour un Etat moderne*, Rapport au Commissaire au plan, janvier.

Thoenig, J.C. (1971), « Le PPBS et l'administration publique. Au-delà du changement technique » *Annales internationales de la fonction publique*, Paris, IIAP, p. 97-114.

Thoenig J.-C. (1998), « Evaluation de l'action publique locale, synthèse », in Kessler, Lascoumes et al., p 313-329

Thoenig J.-C. (1999), « L'évaluation, sources de connaissances applicables aux réformes de la gestion publique », *Revue française d'administration publique*, n° 92, 1999, p. 684.

Thoenig J.-C. (2000), « L'évaluation est-elle une opportunité mal comprise ? », Société française de l'évaluation, Rennes, 15 juin.

Thoenig, J.-C. (2005), « L'évaluation : un cycle de vie à la française » in Lacasse F., Verrier P.-E. (dir.), *30 ans de réforme de l'Etat. Expériences françaises et étrangères. Stratégies et bilans*, Paris, Dunod, pp. 117-127

Tourmen C. (2006), « La mesure en éducation, un outil au service du politique », *L'orientation scolaire et professionnelle*, 35, n°4, p. 315-366.

- Tourmen C. (2007), « Les compétences des évaluateurs. Le cas des évaluateurs de politiques publiques ». Thèse de doctorat, Université Pierre Mendès-France, Grenoble II.
- Van Damme St., (2005), « Les sciences humaines à l'épreuve de la ville : les enjeux d'une archéologie des savoirs urbains (XVII-XXème siècles) », *La Revue d'histoire des sciences humaines*, Auxerre, Sciences Humaines Editions, n°12, pp. 3-15.
- Varone F., Jacob S. (2004), « Institutionnalisation de l'évaluation et Nouvelle Gestion Publique : un état des lieux comparatif », *Revue internationale de politique comparée*, 11(2) p. 271-292.
- Vidaillet B. (2013) *Evaluez-moi ! Evaluation au travail : les ressorts d'une fascination*, Paris, Broché.
- Viveret P. (1989), *L'évaluation des politiques et des actions publiques*, La Documentation française
- Warin Ph. (1996), « L'évaluation : vecteur d'autonomisation des régions », *Pouvoirs locaux*, n°31 IV, p.33
- Wildavsky A. (1972), "The self-evaluating organization", *Public Administration Review*, sept.-oct.
- Wildavsky A. (1973), *Implementation*, University of California Press; Berkeley.
- Zarka Y.-C. et al (2009), « L'idéologie de l'évaluation », *Cités*, Janvier

