

L'ÉVALUATION DES POLITIQUES PUBLIQUES

par **Sylvie Trosa**

Les Notes de benchmarking international

Novembre 2003

Institut de
—
l'entreprise

© Institut de l'Entreprise, 2003
Tous droits de reproduction, de traduction, d'adaptation et d'exécution
réservés pour tous les pays

Directeur de la publication : Jean-Pierre Boisivon, Délégué général de l'Institut de l'Entreprise

SOMMAIRE

PRÉFACE	5
INTRODUCTION	9
1. L'ÉVALUATION ET SES AVATARS : LA LENTE MATURATION DU CONCEPT	11
2. LA DIFFICILE DIFFUSION DES PRATIQUES D'ÉVALUATION EN FRANCE	15
3. MODÈLE FRANÇAIS VS MODÈLE ANGLO-SAXON ? ..	25
4. LES TENDANCES ACTUELLES AU RAPPROCHEMENT DES MODÈLES	33
5. LES FACTEURS CLÉS DE SUCCÈS DE L'ÉVALUATION	41
CONCLUSION	53
ANNEXE 1 : ENQUÊTE DE BENCHMARKING INTERNATIONAL	57
A/ CANADA	58
B/ ALLEMAGNE	59
C/ GRANDE-BRETAGNE	60
D/ ÉTATS-UNIS	61
E/ DANEMARK	62
ANNEXE 2 : LES CONDITIONS DE SUCCÈS DE L'ÉVALUATION	65
ANNEXE 3 : L'EXEMPLE DE L'ÉVALUATION DES SERVICES DE SANTÉ DANS L'ONTARIO	67
SÉLECTION BIBLIOGRAPHIQUE	73
L'AUTEUR	75

L'ÉVALUATION DES POLITIQUES PUBLIQUES

PRÉFACE

L'évaluation des politiques publiques est une discipline d'importance décisive. *Discipline* au double sens du mot : à la fois une véritable technique professionnelle et une exigence qu'une démocratie moderne doit s'imposer.

Les mots sont parfois trompeurs. Pour les non-initiés, l'évaluation désigne simplement une opération de mesure. La réalité est plus complexe. Dès lors que l'on cherche à mesurer les *résultats* d'une politique, il faut être capable de repérer, dans la situation que l'on observe, ce qui est imputable à cette politique et ce qui ne l'est pas. En d'autres termes, toute démarche d'évaluation doit être non seulement une mesure, mais une *explication*. L'évaluation d'une politique publique ne mérite tout à fait ce nom que si elle cherche sérieusement à expliquer quels facteurs ont produits quels effets.

Comme le montre très clairement Sylvie Trosa, il y a une différence majeure entre un système public dans lequel tous les projets d'importance, toutes les politiques définies et outillées, tous les *programmes* au sens de la récente Loi Organique relative aux Lois de Finances (LOLF) donnent lieu à évaluation, initiée si possible *ex ante*, puis répétée de manière périodique, et un système public dans lequel cette discipline, sans être inconnue, n'est pratiquée que de manière épisodique et non réellement intégrée à l'action. Dans le premier cas, l'action publique se donne un maximum de chance de s'ancrer solidement dans les faits ; dans le second, elle court en permanence le risque de se satisfaire de bonnes intentions. Dans le premier cas, elle se place dans une logique de recherche continue d'amélioration, en même temps qu'elle fait place, en cas de nécessité, à des remises en cause fondamentales ; dans le second, elle se prive d'un stimulant essentiel. Dans le premier cas, enfin, elle se dote d'un moyen très important pour concilier au mieux démocratie et efficacité moderne ; dans le second, elle s'expose à méconnaître l'une et l'autre.

Dès lors, quelle place pour l'évaluation des politiques publiques ? Que peut-t-on exactement attendre d'elle ? Comment l'organiser ? A quelles règles doit-elle obéir ? Et sur ces différents plans, comment nous situons-nous par rapport aux pays auxquels nous devons nous comparer ? Autant de questions d'autant plus

L'ÉVALUATION DES POLITIQUES PUBLIQUES

importantes que la crise financière que traverse notre Etat et les difficultés que rencontrent depuis un certain nombre d'années plusieurs de nos grandes politiques publiques ne nous laissent plus de répit.

Nul n'est plus qualifié que Sylvie Trosa pour répondre à ces questions essentielles. L'évaluation des politiques publiques n'a pas vraiment de secret pour elle. Elle l'a pratiquée en France, comme chef du service de l'évaluation du ministère de l'Équipement, puis comme secrétaire générale adjointe du Conseil Scientifique de l'Évaluation. Mais là ne se borne pas son expérience, puisqu'auparavant, elle a été responsable de la politique d'évaluation en Australie, où précisément cette pratique est particulièrement développée. Plus généralement, ce haut fonctionnaire français a une expérience exceptionnelle du management public chez plusieurs de nos voisins, et particulièrement dans les pays anglo-saxons, puisqu'elle a exercé de réelles responsabilités tant à la Banque Mondiale que dans l'administration britannique, puis australienne.

Cette richesse d'expérience lui permet de porter sur les différentes formes que peut prendre l'évaluation des politiques publiques et sur les usages qui peuvent en être faits un regard à la fois très précis et suffisamment distant. Elle montre fort bien que le même concept d'évaluation recouvre des pratiques diverses, qui ont évolué progressivement et correspondent à des philosophies différentes, selon que la priorité est donnée aux finalités de la connaissance ou à celles de l'action, que le processus se veut indolore ou qu'il prend le risque de déplaire, et qu'il peut déboucher ou non sur de vraies remises en cause.

En même temps, elle connaît trop bien ces différentes pratiques, leurs avantages et leurs difficultés pour céder à la tentation simplificatrice de nous présenter telle d'entre elles comme parfaite et telle autre comme condamnable. En particulier, les démarches d'évaluation doivent à la fois être suffisamment indépendantes pour être fiables et en même temps garantir une proximité suffisante entre les principales parties prenantes pour être crédible à leurs yeux. Aussi fait-elle le constat que le « modèle français » et le « modèle anglo-saxon », dont elle spécifie clairement les caractéristiques distinctives, tendent quelque peu à se rapprocher.

Il reste cependant une certitude : par rapport à des pays tels que le Royaume-Uni, l'Australie ou les Etats-Unis, nous sommes franchement en retard dans l'usage que nous faisons de l'évaluation des politiques publiques. Il suffit de prendre connaissance des résultats de l'enquête de benchmarking, qu'on peut lire en annexe de

la présente étude, pour s'en convaincre. La raison de ce retard n'est pas fondamentalement technique, mais d'abord politique et culturelle : à l'inverse des pays anglo-saxons, nous donnons la primauté au combat des idées plutôt qu'à l'épreuve des faits. Nous accordons plus d'intérêt à la définition des politiques qu'à leur mise en œuvre. Cela s'observe aussi bien parmi nombre de parlementaires que de ministres, de hauts fonctionnaires ou encore de responsables syndicaux. Et, disons-le, en chacun de nous.

C'est cela qu'il nous faut changer, et d'urgence. Il ne s'agit pas de renoncer à notre passion pour les idées, mais de mieux la marier avec la discipline des faits. A cet égard, l'excellente étude de Sylvie Trosa vient au bon moment. Je souhaite qu'un maximum de responsables publics lui accordent toute l'attention qu'elle mérite.

Yves Cannac

**Président de l'Observatoire de la Dépense publique
de l'Institut de l'entreprise**

L'ÉVALUATION DES POLITIQUES PUBLIQUES

INTRODUCTION

L'évaluation des politiques publiques souffre, en France, de l'image d'une pratique très technique, longue à mettre en œuvre, difficilement accessible aux administrations par leurs seules compétences et livrée de ce fait exclusivement au bon plaisir des spécialistes.

Rien de plus faux et de plus réducteur, pourtant, que cette image : l'évaluation n'est rien d'autre que la production d'analyses aussi rigoureuses que possible sur les causes d'échec ou de réussite des politiques publiques. En ce sens, elle se distingue du contrôle de gestion, qui va moins loin dans l'interprétation des résultats, et s'attache davantage à la régularité des procédures mises en œuvre.

L'augmentation du nombre de policiers diminue-t-elle effectivement la délinquance, ou ne sert-elle qu'à rassurer les citoyens ? Cette question, posée au début des années 70 par la municipalité de New York, a conduit les services d'évaluation de la ville à envisager une multiplicité d'hypothèses, dont certaines allaient à rebours du sens commun : la délinquance peut être plus efficacement combattue par une meilleure organisation des services de police que par la simple augmentation des effectifs ; en outre un trop grand nombre de policiers peut favoriser chez les citoyens un sentiment de déresponsabilisation dont pâtit finalement l'ordre public. C'est cette recherche des causalités qui est au cœur de l'activité d'évaluation.

Ainsi définie, l'évaluation apparaît comme la condition même de la légitimité de l'action administrative : si l'on ne connaît ni les résultats de l'action administrative, ni les facteurs qui ont conduit à ces résultats, comment justifier l'utilisation de l'argent public ? Et comment motiver des fonctionnaires qui n'auraient pas la moindre idée de l'utilité ni de la pertinence de leurs actions ? De fait, cette nécessité de rendre des comptes aux citoyens et de légitimer l'action de l'administration par la connaissance de ses résultats a été le premier moteur du développement de l'évaluation dans la plupart des pays de l'OCDE.

Après avoir décrit l'émergence du concept d'évaluation (I), cette étude se penche sur les facteurs qui expliquent la lente et difficile diffusion des pratiques d'évaluation en France (II). Mettant en évidence l'opposition historique entre

L'ÉVALUATION DES POLITIQUES PUBLIQUES

un modèle français et un modèle anglo-saxon d'évaluation (III), elle analyse la convergence qui s'opère progressivement entre les deux modèles (IV). Enfin, s'appuyant sur une enquête de *benchmarking* international dont les résultats sont fournis en annexes, elle identifie cinq grands facteurs clés de succès dont la mise en œuvre, en France, permettrait de donner à l'évaluation la place centrale qui lui revient dans la conduite des politiques publiques (V).

1 L'ÉVALUATION ET SES AVATARS : LA LENTE MATURATION DU CONCEPT

I.1 UNE PRATIQUE D'ORIGINE ANGLO-SAXONNE

La pratique de l'évaluation est née au début des années 1960 dans les pays anglo-saxons. Cette antériorité s'explique par la conjonction de plusieurs facteurs :

- la prégnance d'une tradition positiviste, qui se manifeste par le souci d'analyser et de légitimer la conduite des politiques publiques à partir de la connaissance de leurs résultats quantitatifs et qualitatifs ;
- l'importance accordée à la mise en œuvre des politiques publiques, qui contraste avec l'attention portée par les pays latins à la définition de ces politiques ;
- une forte pression du Parlement, soucieux de s'impliquer dans l'examen de l'activité des administrations ; très tôt, par exemple, le *General Accounting Office* s'est investi dans l'évaluation des politiques publiques aux Etats-Unis ;
- la volonté, tôt manifestée par les Directions du budget, de se doter d'outils performants d'allocation des moyens.

On notera par ailleurs que l'évaluation a d'autant mieux « pris » qu'elle était mise en œuvre dans des pays « neufs », comme l'Australie, où les réformes ne sont pas entravées par une culture de droits acquis et par une conception rigidifiée du rôle de l'Etat.

On distingue généralement trois phases dans la maturation du concept.

1.2 L'ÉVALUATION « SCIENTIFIQUE » (1960-1970)

La première phase est celle de l'évaluation qui se veut « scientifique », en procédant à cette fin à des expérimentations conduites à la manière de tests en laboratoire.

Ainsi la ville de New York a-t-elle voulu évaluer l'efficacité de sa police en distinguant trois quartiers, le premier avec des effectifs identiques, le deuxième avec des effectifs diminués de 50 %, le troisième avec des effectifs augmentés de 50 %. Surprise : les effets sur la délinquance furent sensiblement identiques dans les trois quartiers. Pourquoi ? Parce que les sociétés ne sont pas des éprouvettes, mais bien des ensembles humains qui réagissent et s'adaptent. Ainsi, dans le quartier affecté par une diminution des effectifs de police, les citoyens se sont organisés eux-mêmes pour pallier cette insuffisance. Dans le quartier qui bénéficiait d'une augmentation des services de police, l'inverse s'est produit, les citoyens sachant qu'ils étaient mieux protégés s'étant désinvestis d'une vigilance ou d'une action communautaire de police qu'ils effectuaient avant les changements.

Nombre d'évaluations « expérimentales » de ce type ont été abandonnées, car, selon un principe proche de celui d'Heisenberg, il s'est avéré que l'existence même de l'évaluation était susceptible de modifier le comportement des évalués. L'évaluation « scientifique » a néanmoins eu le mérite de mettre l'accent *ab initio* sur la nécessité de disposer, pour procéder à l'analyse des politiques publiques, de méthodes rigoureuses et transparentes, permettant de tirer des conclusions justifiées et justifiables – ce qui reste aujourd'hui encore l'essence même de la pratique évaluative.

1.3 L'ÉVALUATION DE PROGRAMMES (1970-1980)

Dans une deuxième étape, l'évaluation s'est différenciée de la recherche pure. L'échec de certaines évaluations, si longues qu'elles en devenaient obsolètes avant même d'être achevées, a fait prendre conscience du fait que l'exercice ne pouvait être purement spéculatif, mais qu'il devait viser la production d'analyses,

certes aussi objectives et approfondies que possible, mais clairement intégrées dans un processus de décision ; autrement dit, que l'évaluation relevait d'un art du savoir-faire, et d'une tension subtile entre les exigences de la décision et celles de la connaissance.

Cette deuxième étape a donc été marquée par la prise en compte des contingences politiques dans l'élaboration même de l'évaluation. Un exemple aidera à prendre la mesure de celles-ci : lors qu'il s'est agi pour la première fois, et conformément à la loi de 1988, d'évaluer le dispositif autour du Revenu Minimum d'Insertion (RMI), une première mouture du texte a mis en évidence qu'une partie des Rmistes pouvaient être réinsérés dans des activités sociales, mais non dans de l'emploi salarié au sens traditionnel du terme. Cette conclusion n'était pas acceptable par le pouvoir politique, qui entendait démontrer que toute personne était ré-insérable dans un emploi salarié classique ; elle a donc été largement censurée dans la version finale du texte.

Cette évolution a donc permis de donner une finalité très concrète à l'évaluation ; sans sombrer pour autant dans l'instrumentation pure, celle-ci est clairement devenue un outil au service de l'amélioration des politiques publiques.

I.4 LA PRISE EN COMPTE DES REPRÉSENTATIONS DANS LA DÉMARCHE ÉVALUATIVE (1980-1990)

Enfin la troisième phase se caractérise par la prise en compte du poids que pèsent les représentations des acteurs dans le processus de réforme des politiques publiques. Concrètement, elle se traduit par des évaluations qui ne sont plus seulement des études, mais qui sont conçues comme un travail de nature participatif au cours duquel un processus de dialogue et d'apprentissage heuristique se développe entre les acteurs de la politique publique, de sorte que représentations et valeurs évoluent à mesure que l'évaluation progresse.

Cette évolution a eu pour corollaire une complexification des méthodes et des outils mis en œuvre, dont le président de la Cour des comptes australienne s'est fait l'écho en des termes éloquents dans son discours inaugural pour l'année 2003.

« L'évaluation peut apparaître comme un procédé purement technique, au même titre que l'audit. En fait, elle soulève de nombreuses questions d'ordre méthodologique et implique des jugements de valeur. A ce titre, l'évaluation est éminemment politique. Ce discours vise à étudier l'importance des hypothèses méthodologiques et des jugements de valeur dans toute évaluation. Il cherche à montrer en quoi des décisions différentes peuvent changer radicalement les résultats finaux, avec des conséquences notables pour la prise de décision.

Je soutiendrai que les jugements de valeurs et les éléments politiques présents dans chaque évaluation doivent être reconnus. En conséquence, il n'est pas possible d'affirmer qu'il existe une unique bonne méthode. Tout dépend des objectifs et des hypothèses qui sous-tendent le processus d'apprentissage. Ces derniers dépendent de jugements de valeur qui reflètent souvent les intérêts de ceux qui conduisent l'évaluation.

Cette observation nous conduit à un dilemme. D'un côté, les évaluations ont une influence réelle sur l'action politique et la vie quotidienne des citoyens. Il est donc important qu'elles soient conduites de la meilleure façon possible. Et pourtant, la définition d'objectifs et l'interprétation de la dynamique d'apprentissage, tous deux forcément inclus dans l'évaluation, sont des procédures complexes, en constante évolution et influencées par des jugements de valeurs. Il n'y a donc pas une bonne méthode, mais plutôt un ensemble de méthodes et d'évaluations qui dépendent des représentations des acteurs. »

Source : www.anao.gov.au

Cette troisième phase a marqué un net progrès par rapport aux deux précédentes : en admettant que toute réforme des politiques publiques était condamnée à l'échec si elle ne prenait pas en considération, en même temps que les performances des administrations, les représentations des acteurs, elle a fait accéder l'évaluation à l'âge de la maturité.

2 LA DIFFICILE DIFFUSION DES PRATIQUES D'ÉVALUATION EN FRANCE

2.1 LA GENÈSE DE L'ÉVALUATION EN FRANCE

Le développement de l'évaluation en France a connu trois grandes étapes.

L'étape « Rationalisation des Choix Budgétaires »

Dans les années 1970, la politique de « Rationalisation des Choix Budgétaires » (RCB) va de pair avec le développement d'une pensée forte sur l'évaluation *a priori* ; cette pensée connaît son expression la plus aboutie en 1986, avec le rapport dit Nioche-Poinsard, qui préconise une approche scientifique de l'analyse des politiques publiques¹. Toutefois ce rapport, remis juste avant l'alternance, n'est pas repris à son compte par la nouvelle majorité, et ses propositions suivent le déclin de la RCB.

Trois années plus tôt, les mêmes Nioche et Poinsard avaient établi, au cours d'un colloque organisé par le ministère des Finances et le ministère de la Fonction publique, un premier bilan de l'évaluation des politiques publiques en France. Leurs conclusions étaient contrastées. D'un côté, ils relevaient la relative abondance de la littérature sur le sujet, recensant pas moins de trois cents études ayant pour objet l'évaluation a posteriori des effets des politiques publiques, réalisées pour la plupart au cours des cinq années précédant leur enquête. De l'autre, ils pointaient un certain nombre de faiblesses graves du dispositif :

1. M. Deleau, J.-P. Nioche, P. Penz et R. Poinsard, *Evaluer les politiques publiques. Méthode, déontologie, organisation*, Paris, La Documentation française, 1986.

L'ÉVALUATION DES POLITIQUES PUBLIQUES

- une faible visibilité de l'évaluation, peu ou pas identifiée en tant qu'activité distincte de la recherche en sciences sociales, de l'audit ou du contrôle de gestion ;
- une méthodologique souvent approximative : les outils mis en œuvre sont généralement rudimentaires et les démarches suivies peu formalisées ; par ailleurs les évaluations ont un caractère plus descriptif que comparatif ou explicatif, et les tentatives pour mesurer scientifiquement l'impact des politiques restent rares ;
- un manque de rigueur et d'objectivité dans la rédaction des rapports d'évaluation : « 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 ; »
- une très imparfaite insertion de l'évaluation dans l'action politique et administrative : « l'évaluation remplit surtout une fonction de connaissance. Son rôle dans les processus de décision et dans le débat public est actuellement faible². »

L'étape « Renouveau du service public »

L'évaluation connaît un nouveau départ avec la remise au Premier ministre Michel Rocard du rapport de Patrick Viveret, en 1989³. Dans l'esprit de son auteur, l'évaluation doit être un ferment de démocratie au sein de la fonction publique, par l'appel à des expertises distinctes de celles des corps d'inspection, ainsi que par l'organisation d'un débat social ouvert et transparent sur les choix de politiques publiques. Mais le rapport Viveret est mal reçu, et maladroitement mis en œuvre. Il soulève de vives oppositions dans la fonction publique, certains fonctionnaires le considérant comme une tentative déguisée de les opposer à la société civile. En outre la volonté d'ouvrir un débat avant tout choix majeur de politique publique se traduit par une obligation « méthodologique » de faire piloter les évaluations par des « instances d'évaluation », supposées

2. Cf J.-P. Nioche et R. Poinard (dir.), *L'Évaluation des politiques publiques*, Economica, 1984.

3. P. Viveret, *L'Évaluation des politiques et des actions publiques*, Paris, La Documentation française, 1989.

regrouper l'ensemble des parties concernées ; or cette obligation, qui n'a été remise en cause qu'en 2003, a eu pour principal effet, non de pacifier les débats ou d'en clarifier les termes, mais de favoriser la constitution de structures lourdes et rigides dont les membres se sont vécus comme des représentants de leurs administrations ou de leur groupe d'origine, et non comme des acteurs d'une prise de conscience collective.

Le rapport Viveret ouvre néanmoins la voie à un processus d'institutionnalisation de l'évaluation, dont le coup d'envoi est donné par la circulaire sur le renouveau du service public signée par Michel Rocard et publiée au Journal Officiel, le 23 février 1989. L'année suivante est mis en place un Conseil Scientifique de l'Évaluation (CSE), chargé tout à la fois de conseiller les évaluateurs et de contrôler la qualité de l'évaluation des politiques publiques interministérielles. Cet organe peine toutefois à trouver sa place : insuffisamment doté – les deux hauts fonctionnaires qui l'animent ne sont pas en mesure, à eux seuls, de suivre les vastes missions qui lui sont confiées –, il souffre aussi du manque d'implication du pouvoir politique, qui ne trace aucune ligne directrice sur le choix des politiques à évaluer, et ne se livre pas au travail pédagogique nécessaire pour diffuser la pratique de l'évaluation à tous les niveaux de l'administration. Les évaluations conduites par le Conseil Scientifique de l'Évaluation sont donc, pour une large part, le fruit d'initiatives individuelles, lancées sans examen préalable de leur pertinence stratégique (l'évaluation est-elle convenablement intégrée au processus de décision ? ses enjeux sont-ils clairement identifiés ?) ni de leur faisabilité (les informations sont-elles disponibles ? le périmètre de l'évaluation est-il réaliste ?).

Le premier rapport du CSE offre néanmoins l'occasion de dresser un nouveau bilan de l'évaluation des politiques publiques en France, cinq ans après le rapport Nioche-Poinsard.

Dans ses conclusions, ce texte relève d'abord l'intérêt croissant que manifestent les administrations pour l'évaluation, avant d'insister sur les disparités selon les domaines de l'action publique : la politique de l'emploi, la lutte contre l'exclusion, la politique de la ville, le système éducatif, la recherche scientifique, l'aide au développement, les aides au secteur productif, les actions de modernisation administrative sont largement investis par l'évaluation, tandis que d'autres aspects de l'action publique (la préven-

tion sanitaire, les politiques de régulation, la politique macro-économique) demeurent complètement à l'écart du mouvement.

Dans l'examen des pratiques évaluatives, les conclusions du rapport du CSE rejoignent largement celles du rapport Nioche-Poinsard : éclatement du champ, hétérogénéité des finalités et des objets de l'évaluation, absence de références méthodologiques communes, faiblesse des transferts d'expérience d'un secteur à un autre, etc. Le CSE insiste fortement sur le caractère trop peu formalisé des démarches d'évaluation : « la cohérence des différents éléments structurels qui caractérisent un projet d'évaluation (finalités, questionnement, dispositif, méthodes) est en général peu explicite. » Ce flou dans la démarche fait en effet disparaître la spécificité de l'évaluation : « pratiquée dans le contexte du contrôle, l'évaluation désigne parfois un simple recueil de constats et d'observations assortis de jugements à caractère intuitif sur les résultats et les objectifs de la politique. Dans le contexte des études-recherches, l'évaluation se réduit trop souvent à une recherche sur certains aspects d'une politique à partir d'un questionnement sans rapport direct avec les préoccupations des acteurs et des décideurs⁴. »

Le Conseil Scientifique de l'Évaluation est remplacé fin 1998 par un Conseil National de l'Évaluation, qui fait une place plus large aux parties prenantes, et notamment aux élus locaux – qui ne participeront toutefois jamais à ses travaux. Ce nouveau Conseil était supposé pratiquer des évaluations plus simples à mettre en œuvre et plus immédiatement exploitables ; mais en novembre 2001, au moment même où il réussissait à atteindre ces objectifs, par l'allègement des obligations méthodologiques imposées aux administrations et par un meilleur calage de ses actions sur l'agenda gouvernemental, il ne fut pas reconduit, sans raison affichée.

4. Bilan établi par Sylvie Trosa et Bernard Perret in *L'Évaluation de l'expertise à la responsabilité*, Rapport annuel du Conseil Scientifique de l'Évaluation, Paris, La Documentation française, 1991. Quatre ans plus tard, c'est au tour du Commissariat Général au Plan de se livrer, sous la plume de Nicolas Tenzer, à un bilan de l'évaluation des politiques publiques en France. Le Plan met cette fois l'accent sur le faible impact de l'évaluation sur la décision publique : « l'évaluation étant souvent considérée – à tort – comme une forme particulière d'étude, les effets sur la décision publique sont souvent indirects. Rares sont les cas où une évaluation est décidée et menée à bien parce qu'on envisage de modifier une politique publique ».

Paradoxalement, en termes de résultats, le bilan de ces deux instances n'est pas négatif. De 1990 à 1998, le Conseil Scientifique de l'Evaluation a réalisé treize évaluations ; quant au Conseil National de l'Evaluation, il a lancé en 1999 un programme de quinze évaluations, dont la moitié ont été conduites jusqu'à leur terme. Ces chiffres doivent être appréciés à la lumière des exigences inhérentes à la mise en œuvre d'évaluations interministérielles : l'expérience prouve que six mois à un an sont nécessaires dans chaque cas pour cadrer les enjeux de la mission, définir les relations entre les différents ministères impliqués, élaborer les appareils méthodologiques adéquats, identifier les personnalités disponibles pour piloter les actions et composer les équipes capables de les conduire.

L'étape « LOLF »

Au moment même où le Conseil National de l'Evaluation est mis en sommeil, le Parlement vote à l'unanimité, le 2 août 2001, un texte majeur, qui marque une nouvelle étape dans l'histoire de l'évaluation : la loi organique relative aux lois de finances. Ce texte, qui se substitue à l'ordonnance organique du 2 janvier 1959, prévoit la mise en place d'une gestion par programmes. C'est là une évolution de toute première importance : ce qui distingue l'administration par programmes de l'administration traditionnelle, c'est qu'elle est une administration de projets, dont la pérennité n'est pas acquise par sa seule existence mais par l'évaluation de ses résultats. Qui dit programme, dit évaluation, car les actions de l'administration deviennent « contestables » (au sens de la *contestability*), c'est-à-dire qu'elles ne se justifient plus par la seule décision de l'administration, mais par les effets positifs ou négatifs qu'elles produisent sur la société, les citoyens ou l'économie.

Cette avancée majeure sur le plan législatif semble toutefois connaître aujourd'hui de sérieuses difficultés de mise en œuvre. Au niveau des ministères, il n'est plus question d'évaluation mais seulement de contrôle de gestion, ce qui est contraire à l'esprit même des programmes. Formidable levier de modernisation de l'action publique, la LOLF semble s'enliser sous le poids des conservatismes et des corporatismes.

2.2 LES OBSTACLES À LA DIFFUSION D'UNE CULTURE DE L'ÉVALUATION

L'héritage d'une culture administrative tournée vers le contrôle

Pour résumer l'histoire de l'évaluation en France, on est tenté d'user d'une formule provocatrice : on en a beaucoup parlé, on en a fait un peu et pourtant ça n'a jamais vraiment pris. La prégnance d'une culture administrative nationale qui, historiquement, a privilégié le contrôle « punitif » au détriment du diagnostic constructif, a sa part de responsabilité dans cet échec : l'évaluation ne peut se développer dans une culture du blâme. L'évaluation, si elle se borne à repérer des failles, à identifier des coupables et à distribuer des mauvais points, est un exercice facile, rapide mais qui ne contribue que médiocrement à l'amélioration des pratiques de l'administration.

La volonté politique a fait défaut pour bousculer cet héritage culturel. L'évaluation a certes fait l'objet de deux décrets en Conseil des Ministres, en 1990 et 1998, mais ces textes n'ont été accompagnés d'aucun plan sérieux de mise en œuvre, si bien qu'on en est resté aux déclarations de principe. Aucun des gouvernements qui se sont succédé ne s'est réellement impliqué dans le dossier, comme si tous considéraient que la modernisation du service public devait rester une affaire de gestion interne à l'administration, pilotée par les seuls fonctionnaires directement concernés.

Plus profondément, cette pusillanimité s'explique par le potentiel de déstabilisation que recèle l'évaluation. L'expérience prouve qu'une évaluation, si elle veut être efficace, ne peut jamais être totalement consensuelle : même lorsque son projet est élaboré de façon participative avec les évalués, même lorsqu'elle est conduite en impliquant au maximum les acteurs concernés, il est rare que ses conclusions n'aboutissent pas à la remise en cause de situations acquises, d'idées reçues ou de financements existants. Ces remises en cause, bien sûr, peuvent être gérées avec doigté, sur le plan humain comme sur le plan social, et elles peuvent même être jugées extrêmement motivantes par des fonctionnaires attachés à l'utilité de leur action. Mais elles peuvent également déranger des fonctionnaires qui confondent leur mission et la pérennité de leur situation présente, ou des responsables politiques qui abusent des méthodes du marketing pour promouvoir leur action. Cette question n'est pas insoluble mais elle ne doit pas être contournée.

On a sans doute eu trop tendance, pour convaincre du bien-fondé de l'évaluation, à en faire un tableau idyllique d'acteurs qui, dans un dialogue fructueux, apprennent naturellement à se parler et à s'entendre. Cette harmonie peut exister, mais le plus souvent c'est le conflit qui l'emporte. Si certains pays, comme les Etats-Unis ou les Etats du Nord de l'Europe, ont une véritable culture du débat public, d'autres comme la France, le Royaume-Uni ou l'Allemagne s'accommodent mal d'une gestion publique, sereine et sans incrimination, des divergences.

Plusieurs évaluations conduites en France ont ainsi été abandonnées par crainte de déboucher sur des résultats trop déplaisants pour certaines des parties concernées ; c'est notamment le cas d'une évaluation de la politique sportive, dont le volet consacré au financement s'annonçait délicat, ou d'une évaluation consacrée au pilotage de l'offre de formation professionnelle, qui a buté sur les questions de l'équité et de la transparence dans les critères de choix des organismes.

Les mécanismes d'auto-censure expliquent la retenue que manifestent les évaluateurs dès lors qu'il s'agit de s'interroger sur le bien-fondé des politiques examinées, ou tout simplement sur l'existence de voies d'action alternatives. Ainsi, une évaluation consacrée à la politique de la montagne a fait délibérément l'impasse sur le point de savoir si le dispositif d'aides publiques était bien le plus indiqué pour atteindre les objectifs visés⁵. De même l'évaluation de la politique répressive en matière de sécurité routière, par sa formulation même, présupposait qu'une politique répressive dite « contrôle-sanction » était une politique adéquate, sans trop s'attarder sur d'éventuelles politiques alternatives⁶.

L'absence de leadership dans la conduite de l'évaluation

La France a souffert d'un autre handicap : l'absence d'un groupe social se constituant autour de l'évaluation pour en porter les valeurs. Dans la plupart des pays de l'OCDE, celle-ci a immédiatement trouvé de très efficaces relais, soit dans les Parlements, qui y ont vu un moyen d'encadrer les pouvoirs de

5. G. Bazin, *La Politique de la montagne*, Paris, La Documentation française, 1999.

6. M. Ternier, *Mise en œuvre de la politique nationale de sécurité routière par les systèmes locaux de sanctions*, à paraître.

L'ÉVALUATION DES POLITIQUES PUBLIQUES

l'exécutif, soit dans les Cours des comptes, qui y ont vu un moyen d'élargir leur champ d'activité au-delà du contrôle de la régularité, soit encore dans la société civile, qui s'en est emparée comme d'un outil lui permettant de demander plus efficacement des comptes aux administrations.

Aucun de ces relais n'a réellement fonctionné en France. La faiblesse du Parlement dans les institutions de la V^e République a entravé le développement, en France, de pratiques comparables à celles en cours dans les pays anglo-saxons, où les fonctionnaires d'encadrement passent en moyenne deux jours par an sous le crible des questions des parlementaires sur les résultats des politiques publiques. Cette pratique constitue un aiguillon important pour l'évaluation, car bien souvent l'administration n'a d'autre choix que d'y recourir pour obtenir les éléments de réponse qui lui font défaut ; elle contribue en outre à nourrir une réelle culture de la responsabilité.

De son côté, la Cour des comptes a longtemps manifesté une certaine réticence vis-à-vis de l'évaluation, se contentant de placer quelques-uns de ses membres à la disposition d'instances d'évaluation de politiques publiques interministérielles. Cette réserve tenait sans doute, pour une part, au poids inhibant qu'exerçait sa mission historique de contrôle de la régularité des actions de l'administration ; culturellement, la Cour des comptes ne s'inscrit pas dans le registre de la proposition constructive. Par ailleurs, son mode traditionnel de fonctionnement, en auto-commandite, était difficilement conciliable avec le nécessaire processus de programmation de l'évaluation. En tout état de cause, son manque d'enthousiasme pour l'évaluation, alors qu'elle avait vocation à en être le moteur, comme ce fut le cas aux Etats-Unis, en Australie, en Suède ou au Royaume-Uni, n'a pas aidé au développement des pratiques évaluatives en France. On notera toutefois qu'une évolution s'est fait jour depuis quelques années, et que la Cour des comptes tend progressivement à intégrer dans ses travaux la problématique de l'évaluation.

L'éclatement du paysage administratif

Il n'existe plus, depuis la mise en sommeil du Conseil National de l'Évaluation, de dispositif interministériel pour l'évaluation. Les projets jaillissent, désordonnés, à l'initiative de l'Union européenne, d'un ou de plusieurs ministères, d'un ou de plusieurs niveaux de collectivités territoriales, voire d'associations. Si l'on veut évaluer l'ensemble d'une politique publique, il faut par conséquent

pouvoir suivre la chaîne de l'ensemble de ces acteurs, ce qui relève de la gageure en l'absence d'un organisme indépendant capable de les faire s'accorder sur le principe et les modalités de l'évaluation. D'autres pays, confrontés au même problème, l'ont résolu par la mise en place de mécanismes contractuels assortis d'indicateurs et d'obligations de résultat. La France apparaît incontestablement en retard dans cette démarche, les collectivités locales s'arc-boutant sur le principe de libre-administration garanti par l'article 72 de la Constitution pour ne s'engager qu'avec d'innombrables précautions dans des partenariats avec l'Etat. On touche là un problème grave : le risque, à terme, est de n'avoir plus de visibilité sur les résultats que pour un domaine résiduel de l'intervention publique, i.e. les crédits gérés directement par les ministères.

Un démarrage sur des sujets excessivement complexes

D'autres causes, de moindre importance, sont à l'origine des difficultés rencontrées dans la diffusion des pratiques d'évaluation en France. Il est probable, en particulier, que l'évaluation a souffert d'avoir été dirigée initialement vers des sujets d'une excessive complexité ; c'était en particulier le cas des évaluations concernant le domaine social, qui posent toujours de très délicats problèmes de causalité dans la mesure où elles font intervenir non seulement les institutions publiques, mais aussi la famille et tout un ensemble de normes sociales. *De facto*, les premières évaluations ont donc été, pour la plupart, longues, très nuancées dans leurs conclusions et assez peu opérationnelles. Ce constat a alimenté, dans certaines franges de la haute fonction publique, l'idée selon laquelle l'évaluation était par nature un exercice long, coûteux et faiblement productif. En réalité, ces difficultés étaient directement liées aux thèmes retenus : si l'on avait choisi d'évaluer une politique de subventions, l'effet de certaines aides sociales, ou tout autre sujet posant des problèmes de causalité moins complexes, il est hors de doute que l'on aurait obtenu des résultats plus rapides et plus immédiatement exploitables.

Des débats scientifiques qui ont brouillé la perception des enjeux

Avant même de s'y livrer, on a beaucoup débattu, en France, des finalités de l'évaluation ; à certains égards, d'ailleurs, l'affaire n'est pas close. S'il est clair qu'une définition de la notion s'imposait, pour la distinguer en particulier du

L'ÉVALUATION DES POLITIQUES PUBLIQUES

contrôle de gestion, ces débats interminables, de nature essentiellement théorique, étaient-ils réellement nécessaires ?

Lorsque le Conseil National de l'Évaluation publia en 1999 son premier rapport annuel, L'Évaluation au service de l'avenir, il fut vertement critiqué au nom du fait que l'évaluation devait s'attacher à juger le passé, et non s'employer à tirer les leçons du passé pour mieux construire l'avenir : autrement dit, jugeait-on, le rôle des évaluateurs n'est pas d'aider les services publics à se forger des outils leur permettant d'accroître l'efficacité de leur action. Mieux vaut parfois rester dans les limbes du Walhalla que de courir le risque d'apparaître comme un réformiste pervers, est-on tenté de conclure...

Ces débats n'étaient pas sans doute inutiles, serait-ce pour conjurer le risque d'une dérive utilitaire de l'évaluation ; mais dans une phase de mise en place, leur excessive technicité a contribué à brouiller le message simple qu'il fallait délivrer pour convaincre les sceptiques.

3 MODÈLE FRANÇAIS VS MODÈLE ANGLO-SAXON ?

Historiquement, les pratiques de l'évaluation ont suivi, en France et dans les pays anglo-saxons, des voies opposées. Plutôt orienté vers la recherche, le modèle français d'évaluation était peu sensible aux considérations politiques, considérant qu'elles étaient de nature à porter atteinte à son indépendance ; à l'inverse, le modèle anglo-saxon d'évaluation a été immédiatement utilitaire, mesurant les résultats des actions publiques par rapport aux objectifs assignés par l'autorité politique, au risque de perdre de vue la réflexion sur la pertinence de ces objectifs.

3.1 ÉVALUATION MANAGÉRIALE ET ÉVALUATION PARTICIPATIVE

Atravers des instances conçues comme de mini-représentations de la société civile, le modèle français d'évaluation s'est voulu un creuset de la démocratie. Cette ambition s'est clairement définie en opposition à l'évaluation managériale telle que la pratiquent les pays anglo-saxons.

Chacune de ces deux approches présente pourtant des atouts qui lui sont propres.

Parce qu'elle fait intervenir un grand nombre d'acteurs aux intérêts hétérogènes, la force de l'évaluation participative est de produire une délibération réellement démocratique sur le sens des politiques publiques. Sa faiblesse est que poussée à sa limite, elle conduit au relativisme scientifique complet, les résultats de l'évaluation finissant par revêtir moins d'importance que les processus de dialogue et d'apprentissage que celle-ci met en œuvre...

L'ÉVALUATION DES POLITIQUES PUBLIQUES

A l'inverse, la force de l'évaluation managériale tient à ce qu'en se plaçant d'emblée dans une optique opérationnelle, ses recommandations ont plus de chances d'être entendues et mises en œuvre. Son pilotage par la seule administration jette toutefois un doute sur sa capacité à enquêter sérieusement et *a fortiori* à remettre en cause les finalités, les procédures et les outils des politiques publiques.

L'opposition évaluation managériale / évaluation démocratique repose en réalité sur deux postulats. Le premier est que l'évaluation managériale inhibe la capacité de l'administration à interroger ses modes de fonctionnement et à se remettre en cause ; le second, que les évaluations dites participatives permettent de lever toutes les formes de censure.

Or aucun de ces deux postulats n'apparaît totalement convaincant : on peut fort bien plaider qu'une administration réellement motivée sera d'autant plus en mesure de porter le fer dans la plaie qu'elle aura les coudées franches et n'aura pas à ménager les intérêts contradictoires des uns et des autres ; tout comme on peut soutenir que les évaluations participatives peuvent conduire à une neutralisation mutuelle des acteurs, qui préféreront s'entendre sur le *statu quo* plutôt que d'ouvrir la boîte de Pandore.

Une évaluation des aides à la réhabilitation des logements sociaux, réalisée en 1993 à l'initiative du Conseil Scientifique de l'Evaluation, illustre ainsi les limites de l'approche participative telle que la France l'a mise en œuvre : l'évaluation a produit des propositions intéressantes pour une mise en œuvre plus efficace de ces aides, mais n'a pas interrogé leur pertinence, alors que personne n'ignorait que ces aides étaient détournées de leur objet, la réhabilitation, pour servir à l'entretien courant des logements⁷. La crainte de lâcher la proie pour l'ombre – autrement dit, de perdre les ressources existantes sans rien gagner en retour – a figé un consensus parmi les différents acteurs concernés. D'autres dérives du même type mais portant sur des enjeux financiers bien plus conséquents sont à déplorer : des évaluations consacrées aux aides à l'emploi ou aux aides européennes, constituant à l'évidence des rentes de situation plutôt que des leviers de croissance, ont ainsi abouti à des impasses parce que les évaluateurs s'étaient préalablement mis d'accord pour vider leurs recommandations de toute substance.

7. R. Rossi, *La Réhabilitation de l'habitat social*, Paris, La Documentation française, 1993.

Dans les deux cas, les arguments renvoient à une représentation très schématique du comportement des acteurs : d'un côté un « patron », qui par définition n'accepterait pas les remises en cause, de l'autre un collectif qui, dès lors qu'il est invité à s'exprimer, parlerait spontanément le langage de la vérité. Malgré ses aspects caricaturaux, cette représentation continue de peser sur la diffusion de mécanismes d'évaluation en France, où l'on nourrit toujours une méfiance presque irrationnelle à l'égard des commanditaires des évaluations.

3.2 ÉVALUATION INDÉPENDANTE ET ÉVALUATION STRATÉGIQUE

Deuxième opposition fondatrice de l'antagonisme entre modèle français et modèle anglo-saxon d'évaluation, l'opposition entre évaluation indépendante et évaluation stratégique. L'idée sévit toujours, en France, selon laquelle l'évaluation doit être le fait d'évaluateurs indépendants qui n'ont pas à intégrer les paramètres de nature tactique (anticipation des marges de manœuvre de la décision) ou politique (calendrier électoral, etc.). Il y aurait ainsi, en quelque sorte, une incompatibilité entre l'indépendance des analyses et l'insertion dans les mécanismes de la prise de décision.

En réalité, il s'agit là encore d'une opposition bloquante :

– d'une part, il n'est pas possible d'imaginer une approche de l'évaluation qui soit purement stratégique, c'est-à-dire qui se plie à toutes les attentes et contraintes de l'autorité politique, car elle y perdrait toute indépendance et verserait dans ces « évaluations marketing » que certains commanditaires souhaitent pour expliquer « en toute objectivité » que tout va bien ;

– d'autre part, il n'est pas non plus possible de se cantonner à une approche exclusivement scientifique, sauf à accepter la perspective de voir des évaluations « enterrées » parce qu'elles surgissent à un mauvais moment du cycle politique (par exemple juste avant les élections), ou parce qu'elles s'opposent de façon tellement radicale aux positions défendues par le ministre que celui-ci ne peut même les entendre.

Relèvent du même type de travers les évaluations faisant publiquement un bilan très négatif d'une politique publique sans proposer pour autant de mesures visant à corriger les dysfonctionnements identifiés. En France, une évaluation de l'aide aux adolescents en difficulté initiée en 1993 par le Conseil Scientifique de l'Évaluation a ainsi été très mal reçue, car la présentation de tous les défauts de fonctionnement relevés par l'évaluation, qui étaient avérés et dus notamment à la pauvreté des ministères et à l'absence de pilotage politique du dossier, n'a été suivie d'aucune action d'amélioration⁸. L'épisode, on s'en doute, a laissé un souvenir amer aux fonctionnaires. Il en serait probablement allé autrement dans un pays anglo-saxon, où les marges de manœuvre qui peuvent exister pour réparer les éventuelles défaillances constatées (existe-t-il des ressources disponibles ? des possibilités de restructuration ? de formation ? etc...) sont considérées dès avant le lancement de l'évaluation.

3.3 ÉVALUATION INTERNE ET ÉVALUATION EXTERNE

Troisième opposition, l'opposition entre évaluation interne et évaluation externe. Les Français privilégieraient l'interne par souci de mobilisation des acteurs administratifs et des partenaires sociaux, tandis que les Anglo-Saxons seraient davantage portés sur des études ou des commissions indépendantes afin de ne pas hésiter à poser les questions qui fâchent. De nouveau, on peut s'interroger sur la valeur d'une telle opposition.

Il est clair que la coopération des acteurs « internes » est nécessaire, tant pour recueillir des informations que pour faire évoluer leurs pratiques. Il est tout aussi clair que l'intervention d'acteurs qui n'ont pas d'intérêts directs dans la politique publique en cause est fondamentale pour briser les faux consensus.

Pourquoi, alors, s'est-on plu à opposer les deux approches ? Sans doute tout simplement parce que leur combinaison est extrêmement délicate, requérant l'intervention d'une ou de plusieurs personne(s) capables d'assurer l'interface entre des approches et des paradigmes différents. Pourtant elle n'est pas impossible, pour peu qu'aient été convenablement définies, dès l'origine, les

8. J.-P. Cluzel et A. Thévenet, *L'Insertion des adolescents en difficulté*, Paris, La Documentation française, 1993.

règles déontologiques et les finalités de la collaboration qui s'engage : peut-elle aller jusqu'à la remise en cause de programmes et à la restructuration de services ? Ou bien doit-elle seulement viser l'amélioration continue de la qualité des politiques publiques ?

Lorsque l'évaluation est censée, du point de vue du politique, déboucher sur des remises en cause radicales, les pays anglo-saxons préfèrent user de méthodes chirurgicales, i.e. d'évaluations externes, menées assez rapidement et débouchant sur des décisions claires. La France ne s'est pas encore engagée dans des évaluations de ce type, notamment parce que de telles missions sont traditionnellement confiées aux corps d'inspection.

3.4 ÉVALUATION AU SERVICE DE LA DÉCISION ET ÉVALUATION AU SERVICE DE L'APPRENTISSAGE

A quoi sert l'évaluation ? A cette question, régulièrement avancée par ceux que l'évaluation laisse encore sceptiques, il est difficile d'apporter des réponses autrement que sous la forme d'exemples, dans la mesure où aucun pays n'a réellement mis en place de dispositif de suivi des évaluations. On peut néanmoins distinguer deux grands types de débouchés de l'évaluation : dans le premier cas de figure, qui a été privilégié par les pays anglo-saxons, l'évaluation est placée au service d'un processus de prise de décision, et sert directement à infléchir, voire à supprimer des programmes ; dans le second cas de figure, qui a été privilégié par la France, l'évaluation vaut surtout par l'effet qu'elle produit sur les acteurs, notamment au travers de processus d'apprentissage qui peuvent se développer même en l'absence de décision formelle.

L'évaluation comme outil dans la prise de décision

Dans les pays anglo-saxons, la plupart des grandes réformes de politiques publiques sont initiées sur la base d'évaluations mettant en évidence les dérives ou les lacunes des programmes en cours. Le lien entre l'évaluation et la décision n'est toutefois pas toujours affiché, et il serait injuste de ne mesurer l'effectivité de l'évaluation qu'au nombre de programmes supprimés ou officiellement modifiés : la crainte du conflit politique et la recherche de la paix sociale

L'ÉVALUATION DES POLITIQUES PUBLIQUES

conduisent souvent les gouvernements à faire preuve de la plus grande discrétion lorsqu'ils modifient la mise en œuvre d'un programme ou le ciblage de ses bénéficiaires.

Le programme américain d'aide au logement étudiant, qui était initialement conçu sans condition de ressource et profitait aussi aux étudiants qui bénéficiaient du soutien parental, a ainsi été réorienté, après évaluation, vers les étudiants ne disposant d'aucun soutien parental. De même, le système de formation continue dans les entreprises, qui, comme en France, reposait sur un prélèvement de 1 % sur les salaires, procurait des rentes de situation à certains organismes de formation ; l'évaluation ayant mis en évidence que ces rentes étaient liées pour partie à la gestion du système au niveau fédéral, l'ensemble du dispositif a été régionalisé.

L'Australie est un des pays de l'OCDE qui ont donné le plus d'importance à l'évaluation dans le processus de décision. Entre le budget 1990-1991 et le budget 1994-1995, le pourcentage de mesures budgétaires influencées par des évaluations est ainsi passé de 23 % à 77 %, chiffre toujours valable aujourd'hui. Cette montée en puissance est à mettre sur le compte d'un dispositif sévère qui veut qu'une proposition non assortie d'une évaluation ne parvienne même pas au stade de la conférence budgétaire ; les ministères sont ainsi vivement incités à produire les propositions les mieux argumentées possibles. En outre, les économies budgétaires ont été fondées à 65 % sur les résultats d'évaluations, ce qui a permis de cibler avec plus de pertinence les mesures de restriction décidées par le gouvernement.

L'utilisation de l'évaluation pour redéfinir les politiques publiques et réduire les dépenses : l'exemple des politiques de l'emploi en Australie (1996-1997)

En 1996-1997, le nouveau gouvernement australien est décidé à réduire et à réorganiser la dépense publique. Une attention particulière est accordée au marché du travail, dans la mesure où les différents programmes d'aide à l'emploi coûtent chaque année à l'Etat 3,8 milliards de \$ australiens (environs 2,9 milliards de \$).

La ministre de l'Emploi avait défini comme objectif prioritaire du gouvernement l'aide aux chômeurs de longue durée et à ceux qui risquaient de tomber dans le chômage de longue durée. Cet objectif répondait à la fois à des critères d'équité et d'efficacité en assurant une meilleure adéquation entre offre et demande de travail. Dans le même temps, la ministre cherchait à améliorer le rapport qualité-prix des programmes d'aide à l'emploi dans une période de rigueur budgétaire.

Les résultats d'évaluations australiennes et internationales furent largement utilisés pour guider les choix politiques. La ministre souligna les différences de rentabilité des différentes politiques de l'emploi. Cette rentabilité fut évaluée à l'aide d'un indicateur mesurant, pour chaque programme, le coût net d'un nouvel emploi créé. Cette mesure était calculée à partir de la probabilité qu'un chômeur aidé retrouve un emploi à un horizon de six mois, en comparaison à un groupe témoin qui ne participait à aucun programme spécifique.

Cette évaluation révéla que le programme de subvention aux entreprises JobStart avait un coût net de 4 900 \$ australiens par emploi créé, contre 76 600 \$ pour Jobskills, un dispositif de création directe d'emplois. La ministre nota que « le gouvernement concentrerait désormais ses efforts sur les programmes dont la rentabilité en termes de création d'emploi était démontrée ». En conséquence, le programme JobStart fut conservé, tandis que le dispositif Jobskills fut considérablement allégé et concentré sur les chômeurs les plus désavantagés.

L'économie totale réalisée par le gouvernement suite à cette réforme des programmes d'aide à l'emploi fut d'environ 1,5 milliards de \$ australiens sur deux ans.

Source : Guide de la Banque Mondiale sur l'évaluation, 2000.

L'apprentissage comme vecteur d'apprentissage dans la mise en œuvre des politiques publiques

En France, le lien entre évaluation et décision politique apparaît historiquement bien plus distendu. Cette singularité tient en grande partie au degré d'implication bien moindre de l'autorité politique, à la fois dans le choix des projets d'évaluation et dans le suivi de leur mise en œuvre : l'évaluation s'est largement développée hors du champ politique. Dans ces conditions, l'évaluation a

L'ÉVALUATION DES POLITIQUES PUBLIQUES

surtout servi à favoriser l'apprentissage, au sein de l'administration, de nouvelles pratiques et de nouvelles représentations. Cet apprentissage a parfois permis d'obvier à la décision proprement dite, soit parce que le dialogue engagé entre l'administration et les acteurs sociaux avait fait évoluer les représentations de telle sorte que celle-ci était devenue *in fine* inutile, soit parce que les services évalués avaient compris qu'il n'était pas dans leur intérêt de se voir imposer des changements, mais plutôt de les devancer.

L'évaluation du dispositif d'informatisation de l'administration française, en 1993, a suscité une certaine inquiétude chez l'organisme interministériel en charge du dossier, régulièrement critiqué pour ses pratiques centralisatrices, bureaucratiques et peu attentives aux besoins réels des administrations. L'organisme a devancé l'appel en modifiant ses procédures de travail avant la publication des résultats de l'évaluation, afin de prouver son utilité et de retrouver sa légitimité.

De même, l'évaluation de la politique de préservation de la ressource en eau, conduite en 2001 par le Conseil National de l'Évaluation, a rapidement fait apparaître que les périmètres de protection de l'eau n'étaient pas un outil décisif, que, contrairement aux idées reçues, la qualité de l'eau était assez bonne, sauf dans certaines zones délimitées, et que la France disposant de procédures d'épuration efficaces, le véritable enjeu était plutôt celui de la préservation des nappes profondes. La réorientation de la politique publique a pu se faire par accord des acteurs sur le diagnostic établi par l'évaluation, et les changements juridiques n'ont fait qu'entériner les conclusions avancées par celle-ci.

Un examen des pratiques en cours dans les pays de l'OCDE invite toutefois à nuancer fortement cette opposition entre un modèle d'évaluation au service de la décision et un modèle d'évaluation au service de l'apprentissage. De fait, on comprend bien qu'il ne peut guère être satisfaisant de ne jouer que sur l'un de ces deux aspects : la prise de décision formelle peut, comme le démontre abondamment la sociologie des organisations, n'avoir sur les pratiques qu'une influence marginale, voire nulle ; à l'inverse, ne compter que sur les effets d'apprentissage revient souvent à renoncer, par un aveu d'impuissance implicite, à toute réforme d'ensemble d'une politique dont les fondements peuvent être néanmoins reconnus comme viciés.

4 LES TENDANCES ACTUELLES AU RAPPROCHEMENT DES MODÈLES

La portée de cette opposition entre modèle anglo-saxon et modèle français d'évaluation doit être toutefois être nuancée : dans la pratique, et depuis le début des années 90, les pratiques évaluatives des pays de l'OCDE tendent à converger.

4.1 DES PRATIQUES FRANÇAISES VERS LES PRATIQUES ANGLO-SAXONNES

Du côté français, l'évolution est claire. Le dernier rapport du Conseil National de l'Évaluation prône une évaluation « pragmatique » et affranchie du « méthodologisme »⁹. En termes plus concrets, le texte affiche plusieurs objectifs :

- un raccourcissement du délai des évaluations. Ce but n'est pas hors de portée, puisqu'en général, la moitié du temps d'une évaluation est consacrée au choix des présidents et des membres des instances, ainsi qu'aux procédures administratives d'impression et de diffusion. Les pays anglo-saxons ont résolu ces problèmes en composant les instances d'évaluation avant même que les politiques publiques soient mises en œuvre, et en externalisant l'impression et la diffusion des rapports d'évaluation ;
- la simplification des cahiers des charges, resserrés autour de quelques principes méthodologiques fondamentaux, afin que les évaluations puissent être plus aisément adaptées aux spécificités des administrations ;

9. Y. Cousquer, V. Chanut, *Une évaluation à l'épreuve de son utilité sociale : rapport d'activité 2000-2002*, Paris, La Documentation française, 2003.

L'ÉVALUATION DES POLITIQUES PUBLIQUES

- une plus grande souplesse dans le recours aux instances d'évaluation, qui ne sont plus considérées comme systématiquement nécessaires, et dont on privilégie la capacité de travail collectif sur la représentativité, quitte à ce que la diversité des intérêts soit prise en compte par d'autres moyens (auditions *ad hoc*, sous-groupes de travail, groupes de travail locaux, ...)
- une meilleure prise en considération de la dimension nécessairement stratégique de l'évaluation.

Des progrès notables ont déjà été réalisés dans ces directions ; ainsi une évaluation de l'intermodalité en matière de transports a-t-elle été récemment réalisée dans des délais courts, en pilotage direct par le Commissariat Général du Plan, pour répondre à une échéance gouvernementale précise¹⁰.

En ce sens, on peut dire que la France, tout en réaffirmant son attachement à l'évaluation participative, ne désavoue plus l'évaluation managériale : les paramètres stratégiques sont pris en compte, les contraintes du politique également, et les évaluateurs s'attachent désormais à livrer à leur commanditaire des recommandations immédiatement exploitables.

4.2 DES PRATIQUES ANGLO-SAXONNES VERS LES PRATIQUES FRANÇAISES

A l'inverse, le modèle anglo-saxon d'évaluation a intégré un certain nombre de caractéristiques propres au modèle français.

L'intégration de dispositifs participatifs

L'idée d'une évaluation totalement externe qui ne prenne pas en compte le point de vue de l'évalué et sa propre auto-évaluation n'est définitivement plus de mise : le constat s'est en effet imposé que de telles évaluations « de

¹⁰ M. Matheu, *Politique de transport combiné rail / route*, à paraître.

confrontation » figeaient plus qu'elles ne faisaient évoluer les pratiques administratives. Désormais, dans les Cours des comptes des pays anglo-saxons ou même à la Banque mondiale, tout contrôle (au sens générique du terme) va de pair avec :

– une construction en commun des indicateurs, afin que les normes de référence soient partagées par l'ensemble des parties ;

– une définition en commun des procédures de travail et d'échange d'informations ;

– la construction conjointe de recommandations et de plans de mise en œuvre de ces recommandations. Cette construction conjointe relève toutefois d'une obligation de moyens, non d'une obligation de résultats ; autrement dit, il est attendu de l'évaluateur qu'il se livre au travail de pédagogie nécessaire pour faire partager au plus grand nombre les conclusions de son évaluation, non qu'il fasse abstraction de tous les points de désaccord qui subsistent.

L'intégration de dispositifs participatifs dans l'évaluation se fait avec d'autant plus de facilité que dans la plupart des pays anglo-saxons, la société civile collabore traditionnellement à la mise en œuvre des politiques publiques (prévention en matière de santé, surveillance de voisinage, etc.) ; le problème se pose toutefois du financement de ces dispositifs, dont la pérennité ne paraît pas toujours assurée.

Quoi qu'il en soit, l'extrait qui suit du compte rendu des travaux d'un groupe de travail de l'OCDE réuni à Tokyo en septembre 2000 est révélateur des évolutions en cours ; d'abord parce qu'il témoigne de l'intérêt récent de l'organisation internationale pour l'évaluation, alors que l'OCDE ne s'intéressait, jusque dans les années 90, qu'à la mesure des résultats des politiques publiques ; ensuite parce que l'essentiel de ses réflexions portent, non sur la méthodologie de l'évaluation, mais sur la manière de gérer les relations entre les parties prenantes de l'évaluation.

1. Tous les participants ont reconnu qu'il fallait améliorer le processus de communication lié à l'évaluation pour qu'il puisse y avoir appropriation de ses conclusions.

Mais les différences de culture et de pratiques entre ministères sur le rapport au savoir font qu'il n'y a pas de solution magique.

2. Un enjeu corollaire est de mieux cerner les partenaires de l'évaluation avant qu'elle ne démarre et de trouver des moyens appropriés de les impliquer.

3. En revanche, les points de vue demeurent contrastés en ce qui concerne l'indépendance des évaluateurs. Certains pays considèrent qu'il s'agit d'une condition essentielle pour assurer la crédibilité des conclusions tandis que d'autres craignent qu'elle ne porte atteinte aux partenariats qui peuvent se nouer lors d'une évaluation.

4. Néanmoins tous les pays sont d'accord sur le fait que l'évaluation est la composante majeure de l'obligation de rendre compte mais que cette obligation ne doit pas être dissociée de l'exigence d'apprentissage.

5. Le manque d'incitations à évaluer n'est que le reflet de l'absence d'incitations à apprendre et progresser dans les administrations publiques. C'est donc ce problème plus général qu'il s'agit de résoudre.

6. Malgré les difficultés qui lui sont attachées, le principe de la publicité des évaluations a été accepté par tous.

Source : www.oecd.org

L'intégration de l'évaluation dans l'audit de performance et le contrôle de gestion

Jusque dans les années 1990, les pays anglo-saxons ont privilégié, dans leurs systèmes de management, les systèmes de mesure quantitatifs et les audits de performance à l'évaluation. Cette situation est en train d'évoluer rapidement.

La Cour des comptes en Australie s'est ainsi organisée en deux sections : une section chargée du contrôle de la légalité dans l'usage de l'argent public, et une section chargée de l'audit des performances des ministères. L'audit de performance

se scinde lui-même en deux activités : l'appréciation de la performance du fonctionnement des administrations et l'évaluation des politiques publiques que celles-ci conduisent. Dans le dernier cas, la Cour assume une mission permanente de vigie, puisqu'elle doit être en mesure, à tout instant, de répondre à toute question posée par le Parlement.

En théorie, l'audit de performance vise à « apprécier » une organisation à travers des grilles d'indicateurs, régulièrement actualisées, qui permettent d'estimer la qualité du pilotage de l'administration (y a-t-il un plan de management des risques ? La gestion des ressources humaines est-elle intégrée à la gestion stratégique ? etc...). Ces *check lists* sont très utiles dans la mesure où elles permettent de garder la mémoire de ce que l'on appelle « les bonnes pratiques », et de conférer plus de légitimité aux innovations en cours dans les administrations. Leur inconvénient est d'avoir contribué à la dérive formaliste de certaines évaluations, appelées par dérision les « *frameworky analysis* », qui expliquent avec un luxe de détails le fonctionnement de la structure sans jamais s'attarder sur le pourquoi de leur fonctionnement. C'est en prenant conscience des limites d'une telle approche que les pays anglo-saxons ont complété l'audit de performance par une réelle évaluation des politiques publiques, attachée à la mise en évidence des causalités.

Considérons l'exemple d'un ministère qui, selon les critères retenus par l'audit de performance, est considéré comme très mal géré. Une évaluation pourra mettre en lumière les causes de ces mauvais résultats ; elle pourra faire apparaître, par exemple, que le gouvernement n'attend de ce ministère qu'une forte réactivité sur des tâches d'exécution, et qu'il ne se soucie guère des conseils qu'il pourrait donner sur la définition et la mise en œuvre des politiques publiques. Avec des agents épuisés à force de vivre sous la pression de l'urgence, comment le ministère pourrait-il faire figure de modèle de gestion ? La morale de cette histoire est qu'en fait, qu'il s'agisse d'audit ou d'évaluation, on ne peut faire l'impasse sur les finalités assignées à une organisation, même si ces finalités sont implicites. Une mauvaise performance ne peut être redressée que si on en comprend les causes.

4.3 LES LIMITES À LA CONVERGENCE DES MODÈLES

Les développements qui précèdent tendent à établir qu'il n'y a plus aujourd'hui d'opposition substantielle entre les modèles nationaux d'évaluation. Cette convergence doit beaucoup aux réseaux internationaux qui se sont tissés entre les différents acteurs de l'évaluation dans les pays de l'OCDE. Elle a aussi été fortement encouragée par la Commission européenne, qui a institutionnalisé une obligation d'évaluation pour la gestion des fonds structurels et qui associe, comme dans le processus budgétaire australien, une évaluation *ex ante* (cibles, effets, précédents) à toute proposition de dépense nouvelle. Par ailleurs, les Etats membres de l'Union européenne ont spontanément développé des évaluations des actions co-financées par les fonds communautaires pour garantir le sérieux de leur démarche : l'Europe a ainsi littéralement créé un marché de l'évaluation.

La grande différence, aujourd'hui, est celle qui subsiste entre les pays où l'évaluation se développe grâce à des initiatives individuelles et autonomes, comme en France où elle fleurit surtout dans les collectivités locales, et les pays où l'évaluation, considérée comme un véritable levier de modernisation du service public, est un élément à part entière dans le cycle de gestion et de management publics. Chacune de ces deux approches a ses avantages, et, à dire vrai, elles devraient être complémentaires.

La première approche a été illustrée notamment par les travaux de l'ex-Conseil National de l'Évaluation, qui, avec des fonds limités et malgré l'absence d'orientations gouvernementales explicites, a pu, grâce à la bonne volonté des administrations, conduire des évaluations (sur la politique de l'eau¹¹, le rôle de la répression dans la sécurité routière¹², etc.) qui ont conduit à un infléchissement sensible des pratiques administratives.

La seconde approche est née du constat que les actions de modernisation ne doivent pas être des actions *ad hoc*, mais qu'elles doivent s'inscrire dans un nouveau système cohérent de gestion publique.

11. F. Villey-Desmeserets, D. Ballay, D. Tricard, C. Henry de Villeneuve, *La Politique de préservation de la ressource en eau destinée à la consommation humaine : rapport d'évaluation*, Paris, La Documentation française, 2001.

12. M. Ternier, *Mise en œuvre de la politique nationale de sécurité routière par les systèmes locaux de sanctions*, à paraître.

Ainsi le gouvernement canadien promeut-il un modèle de management par résultats qui suit un cycle clairement jalonné :

- définition des résultats à atteindre, en fonction des objectifs gouvernementaux, de l'attente des citoyens concernés et de l'intérêt général ;*
- intégration de ces résultats dans un plan d'affaires contractualisé avec la tutelle ;*
- mise en place des processus permettant d'atteindre ces résultats ;*
- suivi de la mise en œuvre par des systèmes d'information ;*
- mesure des résultats ;*
- évaluation du pourquoi des résultats ;*
- retour sur les résultats / objectifs à atteindre.*

Cette approche s'est aussi développée après l'échec de ce que l'on peut appeler le management par *outputs* (prestations immédiates) des années 1980-2000. Ce mode de management s'était développé en Nouvelle-Zélande, au Royaume-Uni et aux Etats-Unis, à l'initiative de gouvernements qui, estimant que les fonctionnaires déterminaient trop par eux-mêmes leurs priorités, voire ne produisaient pas assez, ont considéré que la première priorité était de faire le point sur les prestations effectivement fournies par les administrations.

Cette politique n'a pas eu que des effets négatifs, dans la mesure où elle a contraint les administrations à sortir d'une vision parfois excessivement abstraite de leurs missions pour se pencher très concrètement sur la qualité des prestations qu'elles rendaient aux citoyens ; elle a par ailleurs permis la mise en place de comptabilités intégrées et de systèmes d'informations performants. Il reste qu'elle s'est traduite par un certain abandon de l'idée même d'évaluation, puisque l'important était de mesurer plus que de comprendre, et une démotivation des fonctionnaires, qui, considérant qu'on ne leur faisait pas confiance, se sont repliés sur une pratique très mécanique de leur travail – les *outputs*, mais rien de plus ni rien de moins. Ce système a fini par se retourner contre les gouvernements qui, faute d'évaluations approfondies

L'ÉVALUATION DES POLITIQUES PUBLIQUES

sur la conduite des politiques publiques, n'ont plus disposé des moyens de répondre de manière satisfaisante aux questions que leur adressaient les parlementaires, et se sont par ailleurs trouvés en manque de propositions venant des fonctionnaires.

Ainsi que l'a reconnu en 1998, lors d'une conférence des sous-directeurs d'administration à Wellington, le ministre de la Fonction publique de Nouvelle-Zélande, cette fixation sur l'output a donné naissance à une génération de fonctionnaires dociles mais effrayés et peu créatifs ; c'est aux ministères qu'incombait, selon lui, la responsabilité de cet échec collectif.

Conscients de ces dérives, la Nouvelle-Zélande, le Royaume-Uni et les Etats-Unis ont renoncé au management par *output* et gèrent aujourd'hui à la fois par *output* et par *outcome*, c'est-à-dire en considérant aussi les effets réels recherchés (par exemple, l'augmentation du taux de placement des chômeurs) ; cette évolution a permis à l'évaluation de retrouver toute sa place.

5 LES FACTEURS CLÉS DE SUCCÈS DE L'ÉVALUATION

L'analyse de la convergence en cours entre modèle anglo-saxon et modèle français d'évaluation a été complétée par une enquête de *benchmarking* réalisée auprès de responsables gouvernementaux et d'experts universitaires de cinq pays (Canada, Allemagne, Royaume-Uni, États-Unis et Danemark), dont on trouvera les résultats en annexes. Ces approches croisées ont permis de mettre en évidence les cinq principaux facteurs clés de succès de l'évaluation :

- la combinaison de contraintes et incitations pour diffuser l'évaluation dans les pratiques administratives ;
- l'établissement d'une véritable planification stratégique de l'évaluation ;
- la mise en œuvre d'une pluralité d'approches dans la pratique évaluative ;
- l'implication et la sensibilisation de l'autorité politique à l'évaluation ;
- la mise en place d'un dispositif d'évaluation professionnalisé.

5.1 LA COMBINAISON DE CONTRAINTES ET D'INCITATIONS POUR DIFFUSER L'ÉVALUATION DANS LES PRATIQUES ADMINISTRATIVES

Il est urgent de considérer l'évaluation comme partie intégrante de l'action administrative. Dans cette perspective, l'institutionnalisation, en France, d'une

L'ÉVALUATION DES POLITIQUES PUBLIQUES

obligation d'évaluation des programmes doit être envisagée, sous la forme d'un mécanisme à deux niveaux :

1. une obligation d'évaluation *ex ante* des propositions budgétaires. Une administration proposant la mise en œuvre d'une politique nouvelle serait priée d'en déterminer les objectifs, si possible chiffrés, les populations cibles, les moyens à mettre en œuvre, et les référents éventuels (*benchmark*) ; sans cette évaluation *a priori*, le ministère des Finances se verrait confier, comme en Australie, le droit de bloquer la proposition dès le niveau des conférences budgétaires. Ce mécanisme qui peut paraître autoritaire s'est révélé en fait plutôt participatif dans les pays où il a été mis en œuvre : mieux informés, les ministères des Finances ont eu tendance à davantage jouer le jeu avec les ministères dépensiers ;

2. une obligation d'évaluer tous les cinq ans au maximum les programmes de la LOLF en cours de mise en œuvre. Une telle obligation impliquerait en particulier de prévoir, pour chaque programme, un financement dédié pour l'évaluation, et la mise en place d'un comité de pilotage de l'évaluation appelé à se réunir durant la phase de mise en œuvre.

En Australie, l'institutionnalisation d'une obligation d'évaluation s'est ainsi traduite par la mise en place d'une semaine de formation gratuite à destination de tous les cadres de catégorie A. La mesure a pu être mise en œuvre en cinq ans.

Au-delà d'un tel dispositif normatif, c'est toute une culture de l'évaluation qu'il s'agit de diffuser au sein de l'administration, notamment au travers d'actions de communication et de formation *ad hoc*.

Cette culture ne pourra toutefois se diffuser que si l'autorité politique se résout à afficher un discours dépourvu de toute ambiguïté sur le sens et la portée de l'évaluation. A cet égard, les remarques qui suivent, extraites du Rapport annuel de la Cour des comptes australienne pour 2002, pourraient être utilement méditées.

« Le service public est confronté à des enjeux de plus en plus complexes qui ne peuvent se traduire par une accumulation de programmes sans évaluation. De sorte que les managers ne peuvent faire l'impasse sur des questions telles que : quels sont nos objectifs ? Comment savons-nous si nous avons réussi ? Les stratégies que nous avons mises en œuvre sont-elles les plus adaptées ? Nos citoyens reçoivent-ils un niveau de service acceptable et sont-ils traités dignement et avec équité ?

Seule l'évaluation couplée au contrôle de gestion peut répondre à ces questions. Mais l'évaluation n'est pas un but en soi : elle sert à améliorer l'action. Une des grandes leçons que nous avons apprises est qu'elle ne doit pas être facultative ou ad hoc mais systématique et toucher tous les programmes. »

5.2 L'ÉTABLISSEMENT D'UNE VÉRITABLE PLANIFICATION STRATÉGIQUE DE L'ÉVALUATION

Le choix des sujets d'évaluation doit faire l'objet d'une véritable planification stratégique, au travers notamment d'une cartographie des acteurs de la politique publique considérée ; cet exercice est un préalable nécessaire pour mieux cerner les attentes de chacun. La prise en compte des attentes des évalués est particulièrement cruciale, dans la mesure où une évaluation brutalement imposée aura toute chance de se voir rejetée par les acteurs sociaux ; mais rien ne serait plus contre-productif que de laisser ceux qui ont intérêt au maintien des situations existantes décider seuls de ce qui sera évalué ou ne le sera pas. Autrement dit, la pratique de l'évaluation suppose une navigation constante et difficile entre le consensus et la fermeté.

Le cas de l'évaluation par la Cour des comptes australienne de l'action publique en matière d'éducation des aborigènes est très révélateur du pouvoir de persuasion dont il faut parfois user pour nouer un dialogue constructif avec les évalués. Le ministère de l'Éducation avait initialement manifesté de très vives réticences, car il savait que le taux d'alphabétisation était très bas, et que la divulgation de cette information lui vaudrait de rudes attaques dans la presse. Pourtant, il a été possible de le convaincre de coopérer avec l'évaluateur, en lui faisant valoir que l'évaluation

lui permettrait justement d'amener le gouvernement, soumis à la pression de l'opinion et des médias, à lui fournir les moyens financiers et juridiques nécessaires pour conduire efficacement son action.

La rénovation du dialogue entre l'évaluateur et les évalués passe par des évolutions dans les méthodes mises en œuvre. Ainsi, dans la plupart des Cours des comptes des pays de l'OCDE, les indicateurs qui sont retenus pour établir l'évaluation sont préalablement discutés avec les évalués ; ce processus participatif contribue à rendre l'évaluation plus incontestable, et, partant, plus efficace.

Dans le même ordre d'idée, certaines Cours des comptes ont fixé à leurs chefs de service un objectif de 90 % de recommandations acceptées par les administrations évaluées. Une interprétation rapide y verrait une incitation au laxisme, l'évaluateur étant incité à se montrer plus conciliant dans ses appréciations pour atteindre son quota. En réalité, l'objectif, souple, des 90 % n'est là que pour mettre l'accent sur la nécessité de convaincre l'évalué de la pertinence des solutions proposées et d'en débattre, étant entendu qu'il demeurera des marges incompressibles de désaccord.

Enfin, la planification stratégique de l'évaluation passe par l'établissement d'une véritable cartographie des acteurs de la politique publique considérée, afin de mieux cerner les attentes de chacun, et par le développement d'une programmation pluriannuelle des évaluations, qui permettra notamment de découper par thèmes et dans le temps certaines évaluations particulièrement complexes.

A titre d'exemple, une évaluation d'un sujet aussi vaste que la politique d'immigration pourrait ainsi être précédée par des évaluations plus limitées portant notamment sur les différentes formes d'entrées (licites et illicites) sur le territoire, les outils dont les ministères disposent pour prévenir l'immigration clandestine, le rôle des négociations diplomatiques, etc.

5.3 LA MISE EN ŒUVRE D'UNE PLURALITÉ D'APPROCHES DANS LA PRATIQUE ÉVALUATIVE

La pertinence d'une évaluation tient pour partie à la pluralité des approches d'investigation qu'elle met en œuvre. Une évaluation purement participative trouve rapidement sa limite dans la difficulté que rencontrent les évaluateurs pour développer de nouvelles approches, non certes par mauvaise volonté, mais simplement parce qu'il leur est pénible de faire le deuil de dispositifs qu'ils ont parfois peiné à mettre en place.

Ce constat conduit à prendre quelque distance avec les propositions du type de celle émise en juin 2003 par le groupe de réflexion sur l'évaluation des politiques publiques locales présidé par M. Gilles Carrez : dans le souci éminemment louable de développer l'évaluation dans les collectivités locales, ce groupe propose la mise en place d'une instance dominée par les élus. Il est à craindre que dans ce type d'initiative, le gain éventuel de légitimité soit plus que compensé par la perte de crédibilité induite.

« Il est proposé la création d'un conseil national des politiques publiques décentralisées qui remplacerait le conseil national des services publics départementaux et communaux. Il serait composé majoritairement d'élus, élus par leurs pairs, afin de lui assurer la plus grande légitimité possible auprès des élus locaux. Il serait présidé par un de ces élus. Il comprendrait en outre des représentants de l'État, par exemple au titre des principales administrations en charge de domaines d'action concernant les compétences des collectivités territoriales et des principaux corps d'inspection et de contrôle. Le mandat des membres du conseil serait d'une durée suffisamment longue pour assurer la continuité des travaux de l'instance. »

De la même manière, s'en tenir à une approche purement externe dans l'évaluation ne serait pas pertinent : cela reviendrait, pour l'évaluateur, à se condamner à ne disposer que d'un accès limité à l'information, et à voir sans cesse remise en cause sa propre légitimité. A l'évidence, le travail heuristique de transformation des pratiques ne se fera pas sans que les principaux acteurs soient impliqués.

Combiner évaluation interne et évaluation externe apparaît donc nécessaire, même si ce travail est extraordinairement complexe : il implique de surmonter les incompréhensions mutuelles, les heurts de perspectives et les conflits de légitimité. A ce titre, il est utile de prévoir la nomination, au sein des comités de pilotage des évaluations, de personnalités « indépendantes » au profil fédérateur, capables de faire travailler ensemble des acteurs dont les langages diffèrent et dont les intérêts mêmes, souvent, divergent.

Concilier le caractère scientifique et le caractère stratégique de l'évaluation relève du même équilibre subtil. Une évaluation qui ne donne pas de gages de rigueur n'est pas crédible et ne sert finalement à rien ; mais une évaluation qui n'aide pas ses protagonistes, par des propositions concrètes, à progresser vers de meilleures pratiques, n'est pas davantage utile. Un « bon » pilotage de l'éva-

L'évaluation doit être considérée comme une action de conduite du changement comme une autre. Le lancement d'une évaluation dont on a toute raison de penser qu'elle va conduire à la remise en cause d'un programme appelle ainsi la définition ex ante d'une stratégie consistant à anticiper ce qu'il adviendra des personnels concernés, à identifier les hommes et les femmes qui seront les mieux à même de piloter le changement, à évaluer les marges de manœuvre disponibles pour financer d'éventuelles actions d'amélioration, à définir les grands axes d'un plan de communication visant à expliquer le pourquoi et le comment de la réforme, etc.

luation associera par conséquent des évaluateurs au profil « scientifique » à des évaluateurs au profil plus « stratégique ».

5.4 L'IMPLICATION ET LA SENSIBILISATION DE L'AUTORITÉ POLITIQUE À L'ÉVALUATION

Parce qu'elle est du savoir inséré dans de la décision, l'évaluation va impliquer des fonctionnaires de différentes cultures et de différents grades, des chercheurs, des consultants, des citoyens, ... bref, plusieurs types d'acteurs qui ne sont pas forcément habitués à travailler ensemble. L'exercice ne pourra être fécond que si tous acceptent de rompre avec le jeu des corporatismes et

de la représentation, et de travailler ensemble, sans récompense évidente, sur des hypothèses et avec des méthodes de travail qui ne leur sont pas spontanément familières ; l'implication du pouvoir politique dans le lancement de l'évaluation, en tant que force de mobilisation représentant l'intérêt général, apparaît donc nécessaire.

La relation de l'évaluation avec le pouvoir politique est toutefois inévitablement complexe : elle ne peut en être détachée, sous peine de compromettre sa légitimité ; mais elle ne peut lui être totalement soumise, sous peine de compromettre son indépendance. Dans les faits, la nature de la relation à l'autorité politique doit être envisagée de façon pragmatique, en fonction de la portée de l'évaluation : de nombreuses évaluations (sur les actions de modernisation de l'administration, sur les dispositifs de formation, etc.) peuvent être menées à bien sans qu'il soit nécessaire de les faire remonter jusqu'au niveau politique. En revanche, les évaluations qui engagent de vrais choix de politiques publiques

Cette conclusion s'impose à la lumière de l'expérience du Conseil Scientifique de l'Evaluation, puis du Conseil National de l'Evaluation. Les deux instances ont souffert d'une procédure de fonctionnement qui sollicitait l'intervention du politique trop en aval du travail de préparation : pourtant intéressé au premier chef, puisque les politiques publiques évaluées étaient par définition interministérielles, le gouvernement ne se prononçait sur l'opportunité des évaluations proposés qu'après de longs mois de travail consacrés à la mise au point de leurs cahiers des charges et de leur méthodologie. Cette pratique, dont le Conseil National de l'Evaluation lui-même a relevé l'absurdité dans son rapport annuel, a conduit à l'abandon d'environ la moitié des évaluations proposées par les ministères... et après beaucoup de temps perdu inutilement.

doivent être soumises très en amont à la connaissance de l'autorité politique, afin que celle-ci puisse les porter et éventuellement les soutenir contre les attaques dont elles pourraient être l'objet.

Dans cette perspective, le rôle du Parlement apparaît aussi décisif. Plutôt que de se focaliser sur l'allocation des ressources *a priori*, qui lui apparaît toujours plus payante sur le plan électoral, celui-ci a tout à gagner à s'inscrire dans l'esprit de la LOLF en demandant des comptes au gouvernement sur la conduite des politiques publiques. La création en février 1999, au sein de la Commission des

L'ÉVALUATION DES POLITIQUES PUBLIQUES

Finances de l'Assemblée Nationale, d'une Mission d'Évaluation et de Contrôle (MEC) qui a déjà entamé un travail actif de réflexion sur les politiques publiques, constitue à l'évidence un premier pas dans la bonne direction, mais le chemin est encore long qui nous sépare des pratiques parlementaires anglo-saxonnes.

L'implication des responsables politiques dans les approches d'évaluation impose de prendre en considération trois paramètres majeurs.

1. Le moment de l'évaluation

La réussite d'une évaluation repose pour une part sur le choix d'un calendrier pertinent : on comprendra sans peine qu'une évaluation qui conduirait à mettre en cause, peu de temps avant les élections, l'action d'un gouvernement sur un terrain sensible, a toutes les chances d'être promptement enterrée... La prise en compte de ce type de considérations est parfois négligée – à tort – par les évaluateurs, plus spontanément familiers des approches scientifiques que des approches tactiques.

2. Le débat public

Les pays où l'évaluation est la moins conflictuelle, comme les pays nordiques, sont aussi les pays où, traditionnellement, les décisions majeures sont précédées de débats publics organisés dans le cadre de commissions réunissant le spectre le plus large des acteurs concernés (cf. aussi la tradition britannique des *Royal commissions*). De fait, l'organisation de débats publics est de nature à faciliter la relation entre l'évaluateur et l'autorité politique : le débat public permet généralement de proposer à celle-ci une gamme plus large de choix ; en outre, il contribue à mieux la protéger, en préparant les esprits aux décisions qui seront finalement annoncées.

3. La question de la responsabilité

De toutes les questions qui se rapportent au mécanisme de l'évaluation, celle de la responsabilité est de loin la plus sensible : dans la culture administrative française, la critique est vécue comme un blâme, lui-même vécu comme un

déshonneur. C'est pour cette raison qu'il importe de réfléchir à la mise au point de mécanismes de médiation permettant de diluer les évaluations, non sur le plan de la causalité, mais sur celui de l'imputabilité. Dans cette perspective, différents procédés peuvent être imaginés : pour l'évaluation des politiques territoriales, ne pas citer nommément le nom des collectivités mises en cause ; pour l'évaluation des politiques nationales, choisir des actions dont les promoteurs ont été assez nombreux pour que l'un d'entre eux ne puisse se sentir directement visé.

De tels contournement ne brillent guère par leur courage, pourrait-on penser. Tout dépend en fait du but visé. Si l'objectif est d'imputer des responsabilités, le contrôle de gestion est probablement plus adéquat. Si l'objectif est de comprendre pour quelles raisons et en fonction de quels critères certaines politiques peuvent être jugées efficaces ou inefficaces, ces ruses dialectiques sont sans doute nécessaires pour éviter les incriminations contre-productives.

5. 5 LA MISE EN PLACE D'UN DISPOSITIF D'ÉVALUATION PROFESSIONNALISÉ

L'un des principaux enjeux, pour cette institution, sera de savoir combiner évaluation scientifique et évaluation stratégique dans l'approche des politiques publiques. C'est d'ailleurs la vocation principale que lui assigne Claude Thélot, conseiller-maître à la Cour des comptes et ancien directeur de la Prospective et de l'Évaluation au ministère de l'Éducation nationale : « la création d'un organisme indépendant d'évaluation permettrait le lien entre recherche et politique, ce que j'appelle la passerelle entre les chercheurs, les évaluateurs ou les experts d'une part, et les administrateurs et les politiques d'autre part. Cette passerelle dans notre pays n'est pas assez construite. Il y a, d'une part, le monde de l'évaluation, de l'expertise, de la recherche qui travaille un peu sur lui-même, et il y a, d'autre part, le monde de l'administration et de la politique qui travaille un peu sur lui-même aussi. Je pense qu'une structure interministérielle devrait avoir pour responsabilité, pour objectif explicite, de bâtir ce que j'appelle « la passerelle », c'est-à-dire de bâtir une capacité de dialogue entre l'expertise et l'action (...) Nous souffrons tout particulièrement dans notre pays d'une insuffisance de compréhension réciproque entre les deux mondes, celui de l'action publique et celui de l'expertise évaluatrice¹³. »

13. Cité in Georges Tron, *Rapport d'information sur les organismes publics d'évaluation et de prospective économiques et sociales*, Assemblée Nationale, n° 876, 21 mai 2003.

L'ÉVALUATION DES POLITIQUES PUBLIQUES

La professionnalisation des pratiques évaluatives en France est un impératif, qui passe notamment par la mise en place, au centre du dispositif, d'une institution chargée à la fois de recueillir les évaluations, d'en élaborer les critères élémentaires et d'en assurer le contrôle qualité. A cet égard, il convient de tirer toutes les leçons de l'expérience du CSE et du CNE, afin de ne pas reproduire les erreurs qui ont compromis leur autorité.

La professionnalisation des pratiques évaluatives passe également par une meilleure prise en compte de l'impact médiatique des études réalisées. Cet enjeu ne doit pas être sous-estimé : nombre d'évaluations dont les résultats étaient balancés, faisant apparaître des points négatifs mais aussi des points positifs, ont vu leurs conclusions détournées par des médias qui n'ont voulu en retenir que les critiques. Ces détournements constituent un puissant facteur d'auto-censure chez les évaluateurs ; de sorte que les évaluations doivent désormais intégrer des réflexions de marketing, afin que leurs résultats ne soient pas présentés de façon unilatérale ou tronquée.

La professionnalisation de l'évaluation implique enfin que l'autorité politique impulse ou *a minima* soutienne les initiatives visant à diffuser une culture de l'évaluation. A cet égard, la France pourrait s'inspirer des actions menées dans les pays anglo-saxons :

- sensibilisation des corps de contrôle à la pratique de l'évaluation, afin de s'affranchir définitivement d'une culture du blâme qui s'attache davantage à la recherche de la faute qu'à celle des solutions ;
- lancement d'un vaste programme de formation des cadres de catégorie A à la logique de programmes et de résultats, et donc à l'évaluation ; l'objectif, à terme, doit être de former, dans chaque ministère ou collectivité publique, des experts du montage et du pilotage d'évaluation ;
- institutionnalisation d'une obligation de rendre compte pour toute entité bénéficiant de ressources publiques, ce qui implique que ces entités disposent d'indicateurs non seulement sur leur activité mais aussi sur leurs résultats, et se livrent chaque année à un *reporting* qu'elles rendent public ;
- mise en place de subventions accordées sous forme contractuelle par la collectivité publique à des associations d'évaluation.

Ce professionnalisme doit toutefois aller constamment de pair avec un solide pragmatisme, qui implique notamment de ne pas développer qu'un seul type d'évaluations – ou, pour le dire autrement, de ne pas mettre tous ses œufs dans

Savoir si la police est plus efficace armée ou désarmée suppose des études comparatives complexes qui demanderont probablement plus d'une année ; en revanche, savoir si le système d'indicateurs d'un ministère est pertinent suppose un groupe de diagnostic avec ses utilisateurs et un regard externe (par exemple un consultant) qui peut travailler en quelques mois. Tous les sujets ne peuvent se résoudre en des évaluations simples, mais a contrario toute évaluation n'implique pas forcément une investigation lourde.

le même panier. L'évaluation doit emprunter plusieurs formes, inscrites dans des horizons temporels allant du très court terme (évaluations « coup de poing ») au moyen / long terme, pour les évaluations les plus complexes. Le projet d'une pratique de l'évaluation « au quotidien » implique de ne pas écarter les évaluations « coups de poing », dès lors que celles-ci s'entourent des garanties de sérieux nécessaires, au premier rang desquelles la prise en compte de la multiplicité des points de vue.

L'ÉVALUATION DES POLITIQUES PUBLIQUES

CONCLUSION

Plus qu'une technique, l'évaluation est une culture, une culture de la « contestabilité », où aucune mesure ne doit jamais être considérée comme acquise pour toujours, et où la légitimité des politiques publiques doit se mesurer à l'aune de leur seule efficacité. Voilà peut-être pourquoi elle a pris de façon plus forte dans les pays anglo-saxons...

Les obstacles...

A quoi peut en effet bien servir l'évaluation dans une société forte de certitudes, qui reconnaît à peine l'existence d'un déficit public et des problèmes qui lui sont liés, où toute régulation repose sur la confiance dans une élite de hauts fonctionnaires, et qui ne ressent par conséquent ni le besoin d'études définissant les problèmes (puisqu'il n'y en a pas) ni le besoin de dessiner des voies alternatives pour les résoudre (puisque le système est bon) ?

Cette question est volontairement provocatrice, mais elle met en lumière trois obstacles dirimants à l'évaluation :

– La méconnaissance des problèmes

Si l'on considère que tout va pour le mieux, qu'il est possible d'augmenter indéfiniment les déficits publics ou que la sortie de la France du concert mondial ne serait après tout pas si dramatique, alors l'évaluation n'est qu'un trouble déplaisant ou tout simplement inutile : *eppur si muove* ...

– La culture du blâme et du secret

Si l'on se satisfait d'un système où toute forme de critique sert d'enjeux à des polémiques, où le débat public est rare et limité à quelques grands spécialistes, alors l'évaluation ne peut que faire peur. Or l'administration demeure dominée

L'ÉVALUATION DES POLITIQUES PUBLIQUES

par une tradition où les réflexions sur les choix de politiques publiques demeurent cantonnées à quelques dizaines de personnes au mieux, et ne sont rendues publiques qu'une fois arbitrées.

– Le syndrome de l'intelligence

La France dispose d'une élite remarquable. Mais chaque force a son revers : cette élite a été formée à la pensée individuelle, dans l'idée que l'intelligence d'un seul ou de quelques-uns résout à elle seule les problèmes ; autant dire que la modestie du « travailleur » d'hypothèses qui doit chercher des preuves et procéder par erreurs-rectifications, tâtonner, tester, faire appel à toutes les intelligences, fussent-elles hors de l'Etat, ne lui est pas encore familière. L'Etat, comme le dirait Michel Crozier, est encore loin d'être modeste...

L'évaluation et ses démons...

Si l'on ne veut éviter que l'évaluation s'enlise dans l'indifférence, il faut combattre ses démons.

Le premier démon est celui de la complexification méthodologique inutile. Certes, certaines évaluations de politiques publiques lourdes, aux causalités diffuses, comme les politiques sociales, seront toujours complexes ; mais l'évaluation peut et doit emprunter différentes formes. L'évaluation consiste d'abord à connaître les résultats des politiques publiques, à s'interroger sur leurs finalités et sur les processus et les jeux d'acteurs qu'elle mettent en œuvre : ce programme apparaît déjà assez ambitieux alors que dans un trop grand nombre de cas, l'administration centrale ne sait même pas si sa politique est mise en œuvre, comment elle l'est, quels en sont les effets et qui en sont les bénéficiaires. Avoir les idées claires sur ces différents points serait déjà un grand pas en avant. L'indifférence au réel est à la limite supportable dans une société riche où les conflits se règlent par l'empilement de nouvelles mesures associées à de nouvelles dépenses, mais la France n'a visiblement plus les moyens de cette désinvolture ; si bien que l'évaluation n'est plus un choix mais une nécessité.

Le deuxième démon est celui de l'appel d'urgence. Trop souvent, l'évaluation n'est lancée qu'en période de crise, lorsque tout va mal – par exemple lorsqu'une administration centrale ne parvient pas à savoir si ses services relaient effecti-

vement sa politique sur le terrain. Or l'évaluation ne saurait faire office de SAMU d'administrations mal gérées, i.e. ne disposant pas de système de remontée de l'information, d'appareil statistique fiable ou d'un contrôle de gestion convenable.

Les opportunités...

Le meilleur moyen de promouvoir l'évaluation serait d'abord de montrer concrètement à l'ensemble des parties prenantes les avantages objectifs que celles-ci peuvent en retirer :

- l'évaluation est utile aux responsables ministériels, car elle leur fournit des arguments objectifs dans le débat sur les choix de politiques publiques et sur les choix budgétaires ;
- elle est utile aux administrations centrales et aux corps d'inspection, car elle affine leur perception de ce qui passe « sur le terrain » ;
- elle est utile aux fonctionnaires, car elle leur suggère de nouvelles pistes pour œuvrer à l'amélioration continue des politiques publiques ;
- elle est utile au contribuable, car elle permet de rationaliser les politiques publiques, d'en supprimer les éventuelles redondances et d'en neutraliser les effets négatifs ;
- elle est utile au Parlement, car elle lui fournit des outils pour mieux contrôler l'action du gouvernement.

Mais tous ces gains, bien réels, n'ont rien d'automatique : l'évaluation exige de savoir se montrer à la fois rigoureux et prêt au compromis ; autrement dit, de marier l'indépendance d'esprit à l'habileté tactique. En toute hypothèse, l'alliance de ces qualités doit être mise au service d'une ambition qui transcende les intérêts particuliers pour œuvrer, avec modestie mais détermination, à l'amélioration continue des politiques publiques.

L'ÉVALUATION DES POLITIQUES PUBLIQUES

ANNEXE I : ENQUÊTE DE BENCHMARKING INTERNATIONAL

A l'appui de cette note, une enquête de *benchmarking* international a été conduite dans cinq pays (Canada, Allemagne, Royaume-Uni, Etats-Unis, Danemark) en vue de mieux identifier les facteurs clés de succès d'une évaluation. Cette enquête a été menée, dans chaque pays, auprès de responsables gouvernementaux et d'experts universitaires. Six grands ensembles de questions leur ont été soumises :

1. Qui est le moteur de l'évaluation : le Parlement, le gouvernement, les ministères individuellement, les Cours des comptes ou une combinaison de ces institutions ?
2. Quel est l'objectif principal poursuivi par l'évaluation : avoir une meilleure connaissance des politiques publiques ? Réduire le coût des actions publiques ? Doter l'administration d'une planification stratégique intégrée et d'un cycle de management cohérent ? Entretenir un processus d'amélioration continue des politiques publiques, fondé sur une base participative ?
3. L'évaluation est-elle affichée comme un réquisit ou bien, facultative, est-elle laissée aux initiatives individuelles ?
4. Quel est le degré d'indépendance des évaluations ? Sont-elles purement internes aux administrations ? Sont-elles rendues publiques ?
5. Y a-t-il une culture d'évaluation, et comment la caractériser ?
6. Y a-t-il un contrôle qualité des évaluations, i.e. une vérification de la fiabilité et de la pertinence des conclusions avancées par les autorités en charge de l'évaluation ? Et qui capitalise les savoir-faire ?

Les résultats de l'enquête sont ici présentés pour chaque pays.

A/ CANADA

Les institutions en pointe dans la diffusion de l'évaluation

Le ministère des Finances et la Cour des comptes ; le Parlement a joué un rôle plus effacé.

Objectif principal de l'évaluation

La production d'études scientifiques sur la pertinence des programmes conduits par l'administration.

Y a-t-il une volonté politique et/ou administrative de promouvoir l'évaluation ?

Oui. Les évaluations font l'objet, au stade préparatoire, de prescriptions très précises, fixées par le ministère des Finances.

Degré d'indépendance de l'évaluation

Les évaluations sont conduites soit en interne soit en externe, mais les cadrage établis par les cahiers des charges sont suffisamment clairs et précis pour que les évaluations soient menées en totale indépendance, n'éluant aucune des questions qui fâchent.

Existe-t-il une culture nationale de l'évaluation ?

Oui. Initiée dans les années 1970, l'évaluation est aujourd'hui complètement intégrée dans la politique de modernisation de l'administration.

Contrôle qualité

Oui. Il est assuré par la Cour des comptes et par le *Treasury Board*.

B/ ALLEMAGNE

Les institutions en pointe dans la diffusion de l'évaluation

En général, le lancement des évaluations relève de la libre initiative des ministères. Le Parlement s'est fait récemment le promoteur de l'évaluation. En revanche, la Cour des comptes et le ministère des Finances se placent encore dans une culture de la régularité juridique plus que de la performance, et constituent donc un frein à l'évaluation.

Objectif principal de l'évaluation

Amélioration continue des politiques publiques, mais sans remise en cause ou questionnement de la pertinence. Il est encore difficile de conduire des évaluations qui questionnent trop directement la pertinence des programmes.

Y a-t-il une volonté politique et/ou administrative de promouvoir l'évaluation ?

Limitée, en dehors des programmes expérimentaux et des cas prévus par la loi ou imposés par Bruxelles (aides européennes), et ceci essentiellement en raison du refus de la Cour des comptes d'aller au-delà du contrôle de la régularité. On remarquera d'ailleurs que les champs les plus couverts sont l'éducation, le logement, le sanitaire et le social, ce qui est caractéristique des pays où l'évaluation est laissée à l'initiative des seuls « convaincus », et non organisée en réquisit du management.

Degré d'indépendance de l'évaluation

Les évaluations sont généralement menées par les fonctionnaires des ministères, éventuellement avec un recours à la sous-traitance pour les études. Peu sont rendues publiques, de peur de soumettre ministres ou hauts fonctionnaires à la critique.

Existe-t-il une culture nationale de l'évaluation ?

Pas réellement : la culture administrative du blâme inhibe le développement de l'évaluation.

Contrôle qualité

Non.

C/ GRANDE-BRETAGNE

Les institutions en pointe dans la diffusion de l'évaluation

Le ministère des Finances a impulsé deux mouvements forts : la rationalisation des choix budgétaires, dans les années 1970, puis, au début des années 1980, le mouvement de déconcentration de l'administration par la création d'agences. La Cour des comptes s'est également placée, depuis vingt ans, dans une optique délibérément constructive du contrôle, n'hésitant pas à l'occasion à dresser des louanges, ou à proposer des solutions aux problèmes identifiés. En revanche, le Parlement a davantage de mal à s'abstraire de la culture du blâme.

Objectif de l'évaluation

Amélioration continue des politiques publiques et responsabilisation des fonctionnaires. L'évaluation est systématiquement intégrée dans le management déconcentré des agences.

Y a-t-il une volonté politique et/ou administrative de promouvoir l'évaluation ?

Oui : l'évaluation est même explicitement désignée comme le principal levier de la politique de modernisation des services publics, et ceci tant par le *Cabinet Office* (services du Premier ministre) que par le *Treasury* ou le Parlement.

Degré d'indépendance de l'évaluation

Elevé dans la phase d'élaboration, plus limité par la suite, en raison du risque d'instrumentation des résultats de l'évaluation dans le cadre d'une polémique publique.

Existe-t-il une culture nationale de l'évaluation ?

Oui. Cette culture est forte mais elle s'attache plus à l'efficacité qu'aux résultats des politiques publiques. Une tendance au changement s'est fait jour récemment, car le gouvernement Blair veut se doter de critères pour hiérarchiser les programmes.

Contrôle qualité

Pas officiellement. Un contrôle qualité est néanmoins réalisé de manière informelle par le *Cabinet Office* et par le *Treasury*, mais il n'est pas rendu public.

D/ ÉTATS-UNIS

Les institutions en pointe dans la diffusion de l'évaluation

A la fois les Finances (*Office of Management and Budget*), héritier d'une longue tradition d'analyse des programmes comme instrument de ré-allocation des ressources, et le Parlement (*General Accounting Office*), investi d'une mission de contrôle et de conseil pour la conduite des politiques publiques. Cette bipolarité a d'ailleurs été une source d'émulation et donc de renforcement de l'évaluation.

Objectif de l'évaluation

Remettre en cause, si nécessaire, la pertinence des programmes.

Y a-t-il une volonté politique et/ou administrative de promouvoir l'évaluation ?

Oui. La politique publique de l'évaluation a été initiée dès le début des années 1960 par le *General Accounting Office*. C'est probablement aux Etats-Unis que l'évaluation connaît son plus haut degré de formalisation.

Degré d'indépendance de l'évaluation

Elevé. Cependant, 20 % seulement des évaluations sont commanditées par le *General Accounting Office* ou par des instances interministérielles indépendantes ; les 80 % restants sont des combinaisons d'évaluations internes et d'études externes commanditées par les chefs d'agences ou de ministères, dont il est difficile d'évaluer l'indépendance.

Existe-t-il une culture nationale de l'évaluation ?

Oui : les Etats-Unis sont marqués par la culture de la *contestability*. L'évaluation est un outil de première importance dans le débat politique : elle est utilisée par l'Etat fédéral, le Parlement, les Etats fédérés et les lobbies pour défendre leurs propres positions ou pour argumenter des choix alternatifs. Les Etats-Unis sont d'ailleurs le pays qui consacre à l'évaluation les moyens les plus importants, puisque le *General Accounting Office* compte près de 1 000 personnes chargées de rassembler l'information et de produire des analyses. Le pays compte par ailleurs plusieurs institutions publiques et privées dédiées à la collecte de données et de statistiques et à la publication d'études : la diversité et la puissance de ces centres d'expertise contribuent à alimenter des débats parfois âpres.

Contrôle qualité

Non, sauf pour les évaluations que le *General Accounting Office* peut examiner.

E/ DANEMARK

Les institutions en pointe dans la diffusion de l'évaluation

L'évaluation se développe tardivement, en raison du mode de régulation des conflits au Danemark, fondé sur l'intensité du débat public, la multiplication des mécanismes de concertation et la recherche permanente du consensus. Toutefois, au début des années 80 et sous la pression budgétaire, le ministère des Finances commence à développer des évaluations d'efficacité ; par la suite, la Cour des comptes se lance dans l'audit de performance, et l'évaluation des programmes sociaux s'accélère avec la création conjointe d'un institut de recherche sociale et d'un ministère du *Welfare*.

Objectif de l'évaluation

Pas d'objectifs partagés entre les différents acteurs de l'évaluation ?

Y a-t-il une volonté politique et/ou administrative de promouvoir l'évaluation ?

Aucune formalisation de politique générale de l'évaluation. Il existe, du fait de la tradition danoise du débat social, une opposition ouverte à des évaluations interrogeant la pertinence des politiques publiques.

Degré d'indépendance de l'évaluation

La question de l'indépendance ne se pose pas réellement eu égard aux objectifs des évaluations, qui visent d'abord l'amélioration continue de l'efficacité des politiques publiques et ne vont pas jusqu'à envisager la mise en cause de celles-ci.

Existe-t-il une culture nationale de l'évaluation ?

Oui et non : la culture dominante est celle du débat social, avant et pendant la mise en œuvre d'une politique publique. Le but recherché est le tissage de consensus entre les groupes sociaux.

Contrôle qualité

Non, pour des raisons qui découlent clairement des analyses précédentes : le contrôle qualité n'apparaît pas essentiel dans la mesure où les évaluations ne s'engagent pas dans des conclusions qui pourraient être ressenties comme conflictuelles.

Cinq conclusions s'imposent à l'issue de cet examen des pratiques par pays :

– La diversité des modèles d'évaluation de par le monde renvoie à la diversité des cultures administratives et politiques nationales ; à cet égard, on peut opposer la tradition danoise du consensus à la tradition américaine du débat ouvert, contradictoire voire conflictuel.

– L'évaluation peine à se développer lorsqu'elle ne joue pas simultanément sur les deux tableaux du travail scientifique (modèle allemand) et du débat social (modèle danois).

– Une pratique efficace de l'évaluation requiert une professionnalisation des savoir-faire, des équipes dédiées et des moyens conséquents pour recueillir et analyser les données ; on rappellera que le *General Accounting Office* compte plus de mille agents.

– L'évaluation peut se développer sans faire l'objet d'une véritable politique publique, mais en restant cantonnée à des analyses au cas par cas, elle compromet son autorité et ne peut généralement assurer un niveau constant de qualité ; on peut opposer à cet égard le groupe que forment le Canada, l'Australie et les États-Unis au groupe que forment l'Allemagne, la France et le Danemark.

– Pour être crédible, l'évaluation doit nécessairement poursuivre plusieurs objectifs : obliger l'administration à mieux rendre compte de son action, améliorer ses pratiques et interroger la pertinence de ses programmes et de ses objectifs. L'exclusion de ce dernier objectif tend à décrédibiliser l'évaluation.

L'ÉVALUATION DES POLITIQUES PUBLIQUES

ANNEXE 2 : LES CONDITIONS DE SUCCÈS DE L'ÉVALUATION

Ces facteurs clés, qui reprennent en les synthétisant les différents points évoqués au chapitre 6, sont indiqués ici par ordre d'importance.

Facteurs clés de succès	Facteurs d'échec
Créer une culture de l'évaluation ; autrement dit, convaincre de la nécessité de disposer d'études, de chiffres et non de sa seule intuition personnelle	Limiter l'évaluation à quelques spécialistes
Une impulsion politique claire, relayée par la haute fonction publique	Utiliser l'évaluation comme un luxe, quand on en a le temps et l'argent, ou à l'inverse n'y recourir qu'en cas de crise grave
Institutionnaliser une obligation d'évaluer les programmes à échéances régulières et budgétiser l'évaluation au moment du vote des programmes	Evaluer en fonction de considérations d'opportunité politique ou de la disponibilité des experts
Une formation généralisée des responsables de programmes à la logique de programmes et donc au pilotage de l'évaluation, qui doit être continu et stratégique	Ecrire un cahier des charges, puis sous-traiter l'évaluation et en attendre les résultats, sans dialogue en cours d'évaluation

L'ÉVALUATION DES POLITIQUES PUBLIQUES

Penser une véritable stratégie de l'évaluation (quand la lancer ? Quel pilotage ?) ; évaluer <i>ex ante</i> les marges de manœuvre disponibles pour un changement	Lancer une évaluation sans savoir pourquoi et en ignorant le jeu d'intérêts qu'elle risque d'éveiller
Bâtir une stratégie de communication autour des conclusions de l'évaluation : comment présenter celles-ci de telle manière qu'elles soient acceptées par le plus grand nombre ?	Se contenter d'une simple publication d'un rapport suivie d'une conférence de presse
Prendre en compte dès le début l'ensemble des parties prenantes : les identifier, les associer et entretenir un dialogue avec eux	Créer une instance d'évaluation qui, sous couvert d'un dialogue entre parties prenantes, soit le théâtre d'une négociation « classique » entre représentants d'intérêts
Susciter un milieu de spécialistes de l'évaluation comme il peut en exister du contrôle de gestion – les apprentissages méthodologiques étant les mêmes dans les deux cas	Laisser perdurer la situation, qui caractérise la France, d'une dramatique insuffisance de l'offre d'évaluateurs par rapport à la demande
Susciter un changement culturel dans les institutions de contrôle en faveur de l'évaluation	Disjoindre les logiques du contrôle et de l'évaluation
Disposer d'un organisme de contrôle qualité de l'évaluation, afin d'apprécier, non leur qualité intrinsèque, mais le caractère fondé et recevable de leurs conclusions	Réaliser ce travail de manière pragmatique et informelle, le laissant ainsi à la compétence aléatoire des individus chargés de tirer des conclusions de l'évaluation

ANNEXE 3 : L'EXEMPLE DE L'ÉVALUATION DES SERVICES DE SANTÉ DANS L'ONTARIO

Ce petit guide méthodologique, établi pour appuyer le programme d'évaluation des services de santé dans l'Etat de l'Ontario (Etats-Unis), offre l'avantage de présenter avec concision mais exhaustivité, en quatre grandes sections (Quand ? Comment ? Qui ? Pourquoi ?), les principes qui doivent guider la pratique de l'évaluation. Pour une approche plus approfondie de la question, on se reportera au guide établi au milieu des années 90 par le Conseil Scientifique de l'Evaluation¹⁴.

ONTARIO HEALTH UNITS

The Guiding Principles for Program Evaluation in Ontario Health Units provide a framework for strengthening the evaluation of public health programs. The Principles outline when, how, and why evaluations should be conducted and who should be involved. Evaluation activities in Ontario health units should be based on the ideals represented in the Principles.

Definitions

A **principle** is defined as a general law which guides action.

14. Conseil Scientifique de l'Evaluation, *Le Petit guide de l'évaluation*, Paris, La Documentation française, 1996.

A **program** is defined as a series of activities supported by a group of resources intended to achieve specific outcomes among particular target groups.

Program evaluation is the systematic collection, analysis and reporting of information about a program to assist in decision-making.

Stakeholders are individuals and groups (both internal and external) who have an interest in the evaluation, that is, they are involved in or affected by the evaluation. Stakeholders may include program staff or volunteers, program participants, other community members, decision-makers, and funding agencies.

Guiding Principles

WHEN

Integrated Program Planning and Evaluation

- Evaluation should be an integral part of program management and should occur during all phases of a program.
- All program plans should include how and when programs will be evaluated.

HOW

Clear Description of the Program

- The program being evaluated should be clearly described, especially process and outcome objectives, as well as intended target groups. Program logic models should be used when appropriate.
- Program objectives that are not specific should be clarified before continuing with further evaluation activity.

Explicit Purpose for Identified Need

- The purpose of any evaluation should be explicit and based on identified decision-making needs.

Specific Evaluation Questions

- Evaluation questions should be specific and clear.
- Evaluation questions should be based on the need to answer key management questions.
- The developmental stage of a program, its complexity and the reason for evaluating should be considered in formulating evaluation questions.
- Evaluation questions directly reflect a program's process and/or outcome objectives.

Ethical Conduct

- Members of the evaluation team should consider the ethical implications of program evaluation to ensure the rights of participants in the evaluation are respected and protected.

Systematic Methods

- The evaluation questions should drive the evaluation methods utilized.
- A review of the literature and a scan of evaluation activity in relevant program areas in other health units should be carried out at the outset of the evaluation.
- New data should not be collected if existing information can adequately answer evaluation questions.

L'ÉVALUATION DES POLITIQUES PUBLIQUES

- The most rigorous evaluation methods should be used given time and resource limitations.
- Evaluation should employ information (quantitative or qualitative or both) gathered from a variety of sources with varying perspectives.

Clear and Accurate Reporting

- Evaluation reports should include a description of the program and its context, the purpose of the evaluation, information sources, methods of data analysis, findings and limitations.
- Evaluation reports should be presented in a clear, complete, accurate, and objective manner.

Timely and Widespread Dissemination

- The dissemination of evaluation findings to stakeholders should be timely.
- Evaluation findings should be shared with other Ontario health units when appropriate.

WHO

Multidisciplinary Team Approach

- The evaluation team should include a variety of people who have adequate knowledge of the program, its participants, and program evaluation.
- Responsibilities should be agreed upon at the beginning of the evaluation. One person should be responsible for the overall management of the evaluation.

- The evaluation team should seek technical advice, support, and/or training, when necessary.
- Members of the evaluation team should continuously work toward improving their program evaluation skills ; team members with evaluation expertise should support this learning.

Stakeholder Involvement

- Stakeholders should be consulted and, if appropriate, involved directly, throughout the evaluation process, within time and resource limitations.
- Stakeholders' interests, expectations, priorities, and commitment to involvement should be assessed at the outset of the evaluation.
- Communication among stakeholders should be honest and open.
- Evaluation should be sensitive to the social and cultural environment of the program and its stakeholders.

WHY

Utilization of Evaluation Findings

- Program managers should formulate an action plan in response to evaluation findings.
- Evaluation findings should be used to support decision-making.

L'ÉVALUATION DES POLITIQUES PUBLIQUES

SÉLECTION BIBLIOGRAPHIQUE

1. Ouvrages

Perret (Bernard), *L'Évaluation des politiques publiques*, Paris, La Découverte, 2002.

2. Rapports

Conseil National de l'Évaluation, *L'Évaluation au service de l'avenir*, Paris, La Documentation française, 1999.

Conseil National de l'Évaluation, *Une évaluation à l'épreuve de son utilité sociale*, Paris, La Documentation française, 2003.

Tron (Georges), *Rapport d'information sur les organismes publics d'évaluation et de prospective économiques et sociales*, Assemblée Nationale, n° 876, 21 mai 2003.

3. Sites Internet

Australie : site de l'*Australian National Audit Office* (www.anao.gov.au)

Banque Mondiale : section « Evaluation » du site (www.worldbank.org/oed)

Canada : site du *Treasury Board* (www.tbs-sct.gc.ca)

Etats-Unis : site du *General Accounting Office* (www.gao.gov)

France : sites du Conseil National de l'Évaluation – Commissariat Général du Plan (www.evaluation.gouv.fr/cgp) et de la Société Française d'Évaluation (www.sfe.asso.fr)

Royaume-Uni : site du *National Audit Office* (www.nao.gov.uk)

L'ÉVALUATION DES POLITIQUES PUBLIQUES

L'AUTEUR

Sylvie Trosa : ancien élève de l'ENS et docteur en droit public, elle a consacré sa carrière à la modernisation du service public et à la réforme de l'Etat. Elle a initié l'évaluation en France et est une des rares fonctionnaires françaises à avoir exercé des responsabilités de haut niveau dans d'autres pays, en particulier au *Treasury* britannique et en tant que chef de service à la Direction du Budget en Australie. Elle a publié six livres, consacrés notamment à la rémunération au mérite dans la fonction publique, à la comparaison des processus de modernisation de l'administration au sein des pays de l'OCDE, et au développement de la démarche contractuelle au sein de l'Etat ; le dernier en date est consacré à la Loi Organique relative aux Lois de Finances (*Le Guide de la gestion par programmes : vers une culture du résultat*, Editions d'Organisation, 2002). Elle est actuellement chargée de mission auprès du directeur du personnel et des finances au ministère de la Santé et des Affaires sociales.