

L'EVALUATION AU SERVICE DE L'AVENIR

RAPPORT ANNUEL 1999

CONSEIL NATIONAL DE L'ÉVALUATION

Avant-propos

Voici plus d'un an que le Conseil national de l'évaluation, installé le 16 février 1999, s'efforce d'accomplir la mission que lui a assignée le décret du 18 novembre 1998 : relancer l'évaluation des politiques publiques et prendre la suite du Conseil scientifique de l'évaluation.

Fort de ses quatorze membres, s'appuyant sur son rapporteur général et disposant du concours du Commissariat général du Plan, qui assure son secrétariat, le CNE a bien engagé sa mission : fin juin 1999 il proposait à la décision du Premier ministre un premier programme d'évaluations. Depuis, les cinq instances d'évaluation décidées en juillet ont été constituées et mises en place grâce aux services du Plan entre novembre 1999 et avril 2000. Elles s'emploient à respecter leur programme dans le meilleur climat : les auditions récentes par le CNE des présidents et rapporteurs de trois de ces instances en témoignent.

Ces premiers accomplissements tiennent à l'engagement motivé et compétent de ses membres, en particulier ceux qui se sont portés rapporteurs sur tel ou tel thème d'évaluation, et à l'articulation trouvée entre eux et Sylvie Trosa, notre ardente rapporteure générale, d'une part, et le Service de l'évaluation et de la

modernisation de l'État qu'anime Nicolas Tenzer, secondé très efficacement par Claire Guignard-Hamon, d'autre part. Les travaux du CNE ont soudé ses membres en équipe, une équipe ouverte sur les pratiques diverses de l'évaluation en France : les liens établis avec la toute nouvelle Société française d'évaluation donnent au CNE une vision réellement nationale de sa tâche.

Cette vision, nourrie d'abord de l'examen des projets d'évaluation qui lui étaient soumis, puis du bilan méthodologique d'évaluations passées, françaises ou étrangères, a étayé l'idée pragmatique de l'évaluation que le CNE s'efforce de faire passer : celle d'une analyse, orientée par les questionnements des commanditaires d'une politique publique, et voulue circonscrite dans la durée, et donc dans son champ, pour pouvoir servir aux décisions d'avenir sur cette politique.

Dans sa démarche qui doit porter sur l'ensemble des politiques publiques, et pas seulement sur celles de l'Etat, le CNE a cependant rencontré une difficulté, s'agissant de la représentation en son sein des trois niveaux de collectivités territoriales – régions, départements, communes – et de la présence des représentants désignés sur proposition de leurs associations. Il est en effet difficile pour des élus chargés d'obligations locales nombreuses d'assumer de surcroît ces charges nationales. Mais ou cette présence devient effective, ou bien une suppléance devra être organisée, en accord avec les associations concernées : il en va du sort de l'élargissement souhaité du champ du CNE par rapport au CSE. Pour sa part, le CNE s'emploiera, dès son programme de travail 2000, à organiser des réunions « décentralisées » pour marquer son intérêt pour les évaluations locales et leur bonne articulation avec les travaux nationaux.

Le CNE dégage un autre enseignement de sa première année de fonctionnement : il est possible de raccourcir substantiellement le délai de constitution des instances d'évaluation. À condition, et le Commissariat général du Plan s'en accorde avec le Conseil, de pouvoir pressentir des présidents et rapporteurs potentiels au moment où le programme des nouvelles évaluations est proposé au Premier ministre pour décision.

Nous nous y emploierons dès les prochains mois en préparant le projet de programme des évaluations pour l'année 2000, en vue d'une décision à l'automne. Pour ce faire, le CNE utilisera l'ensemble des projets déjà formulés par les administrations. Il entendra aussi de grands responsables publics pour confronter leurs visions et priorités en matière d'évaluation et de réformes de

l'Etat à celles déjà échangées entre ses membres. Ces auditions aideront le CNE à accomplir la mission qui lui a été confiée, et à la mieux faire connaître.

C'est aussi l'objet du présent rapport d'activité qui vise l'ensemble des parties prenantes à l'évaluation des politiques publiques et leur permet de mieux situer leur propre pratique. Ce rapport, au-delà d'un compte rendu de l'activité du CNE, rassemble en effet une série de contributions de très haute tenue sur l'état de l'art de l'évaluation dans le monde.

Je tiens enfin à saluer tout particulièrement Sylvie Trosa, rapporteure générale du CNE jusqu'au mois de mai 2000, pour le travail qu'elle a accompli et le rôle d'animation et de facilitation qui fut le sien. Là où elle a choisi d'aller, le CNE sait pouvoir compter sur une amie et sur une pionnière de la réforme de l'Etat et de l'évaluation publique.

Yves Cousquer

Président du CNE

Sommaire

<u><i>Avant-propos</i></u>	<i>1</i>
<u><i>Introduction</i></u>	<i>9</i>
<u><i>Chapitre Premier La nouvelle procédure</i></u>	<i>12</i>
1. L'activité du Conseil en 1999	12
2. L'originalité du nouveau dispositif	14
2.1. Une procédure simplifiée.....	14
2.2. Une étroite collaboration avec le Commissariat général du Plan	15
2.3. Le maintien du contrôle de la qualité des évaluations	15
2.4. Les étapes de la pratique actuelle en 1999	16
<hr/>	
<u><i>Chapitre II Les conclusions méthodologiques tirées des premiers projets</i></u>	<i>18</i>
1. Des conclusions stratégiques	18
1.1. Le temps nécessaire à une évaluation	19
1.2. Des évaluations au service de l'avenir.....	20
1.3. L'évaluation, partie intégrante d'une culture du résultat	21
1.4. L'évaluation est plus qu'un bilan ou un constat, c'est la compréhension des causes des résultats.....	22
1.5. L'évaluation nécessite des instances d'évaluation pluralistes, compétentes et efficaces	24
2. Des conclusions méthodologiques	27
2.1. Où est le problème?	28
2.2. Plus de quantitatif et plus d'écoute.....	31
3. Les pistes à développer	34
3.1. La participation des élus locaux	34
3.2. La culture d'évaluation dans les administrations publiques	35
3.3 Une nécessaire liaison avec les travaux de la Mission d'évaluation et de contrôle du Parlement.....	35
3.4. Une approche plus internationale de l'évaluation.....	36
3.5. Une utilisation de la technologie informatique au service de l'évaluation.....	36

3.6. Recommandations	37
<hr/>	
<u><i>Annexe I Discours du ministre de la Fonction publique, de la Réforme de l'État et de la Décentralisation</i></u>	40
<u><i>Annexe II Lettre du CNE au Premier ministre</i></u>	46
<u><i>Annexe III Evaluations en cours</i></u>	62
<u><i>Contribution I Contributions au séminaire du CNE du 25 novembre 1999</i></u>	63
1. Quelques concepts clés pour définir un projet d'évaluation par Eric Monnier (C3E)	63
1.1. La valeur ajoutée de l'évaluation	63
1.2. Relier les différents types d'évaluation.....	66
1.3. L'évaluation ex-ante.....	67
1.4. L'évaluation à mi-parcours	68
1.5. L'évaluation ex-post.....	69
2. Les cycles de programmation successifs	69
3. Formuler les questions évaluatives.....	71
3.1. Liste des questions possibles.....	71
3.2. Des questions prioritaires.....	72
3.3. Critères d'évaluation	73
4. Répartir les rôles dans le dispositif d'évaluation	74
5. Vers un renforcement de la dimension politique ?	77
6. Panorama des pratiques d'évaluation en Europe par Jacques Toulemonde	78
6.1. L'évaluation, produit d'importation ou initiative nationale ?.....	79
6.2. L'évaluation : exercice administratif, outil de gestion ou devoir démocratique ?	86
6.3. L'évaluation vécue comme une obligation administrative.....	87
6.4. Quelle place pour des modèles nationaux ?	90
6.5. Références bibliographiques	93

Contribution II Les pratiques de l'évaluation française des contrats de Plan État-région 94

1 La superposition des aires de compétences et l'emboîtement des niveaux....95

2. Le développement de l'évaluation et les raisons de ce développement98

***3. Le cas de la mise en œuvre partenariale des politiques publiques : les contrats de Plan État-région en France de 1993 à 1999*..... 101**

3.1. Les pratiques partenariales de l'évaluation des contrats de Plan État-région (1993-1999)..... 101

3.2. Les organisations nécessaires à l'évaluation : instances et dispositifs partenariaux ... 107

4. Réalisation coopérative des évaluation de contrats de Plan État-région (1993-1999) 112

Contribution III Éléments de réflexion issus de l'expérience internationale.... 118

1. Les limites de l'évaluation dans les années 1970-1980 : l'absence d'interministérialité, le blâme et la faible légitimité de l'État..... 119

1.1. De la mesure des performances à l'évaluation 119

1.2. Du contrôle et du blâme à l'évaluation 120

1.3. L'évaluation est aidée mais non résumée par les informations venues du marché 120

2. Approche scientifique et approche participative 121

2.1. L'approche scientifique : un instrument au service de remises en cause radicales..... 122

2.2. L'approche participative : un outil au service de l'amélioration de la mise en œuvre des politiques publiques..... 122

3. La nécessité de générer une demande effective d'évaluation au plus haut niveau..... 124

4. La nécessité de produire des attentes réalistes..... 125

5. Le besoin d'institutionnaliser l'évaluation..... 125

6. Où rattacher l'évaluation ? 125

7. Planifier l'évaluation apparaît comme un réquisit de base 126

8. S’assurer de l’opportunité des questions évaluées est aussi décisif.....	126
9. La publicité des évaluations et des recommandations ne fait pas encore l’objet d’un accord au niveau international	126
10. Assurer la participation de l’ensemble des parties prenantes	126
11. Garantir la qualité méthodologique des évaluations	127
<hr/>	
<i>Contribution IV Évaluation scientifique et évaluation participative.....</i>	<i>128</i>
1. Introduction.....	128
2. Définition de l’évaluation partenariale	130
3. Quels sont les arguments avancés en faveur de l’évaluation partenariale ?	133
4. La critique de l’évaluation partenariale	134
<hr/>	
<i>Contribution V Évaluation et décision budgétaire.....</i>	<i>145</i>
1. Introduction.....	145
2. Origine et étapes du développement	146
2.1 L’origine des réformes du secteur public remonte à 1983	146
2.2. Un intérêt croissant porté à l’évaluation	148
2.3. Une stratégie d’évaluation formelle.....	149
2.4. Évolutions ultérieures	151
3. Le développement de la capacité d’évaluation : les faits	153
3.1. Préparation de l’évaluation	154
3.2. La conduite de l’évaluation.....	155
3.3. Qualité de l’évaluation.....	156
3.4. L’utilisation de l’évaluation	157
3.5. Exemple d’utilisation de l’évaluation pour aider un gouvernement à supprimer et réorienter les priorités de ses programmes	159
4. Les facteurs favorisant ou freinant la réussite : ce qui a fonctionné, les défaillances et qu’elles en sont les raisons.....	160

4.1. Le ministère des Finances (MdF)	160
4.2. Services opérationnels	161
4.3. Autres facteurs	163
5. Conclusions de l'expérience australienne	165

Contribution VI L'évaluation des politiques et des programmes au niveau du gouvernement fédéral canadien..... 167

1 Quelques définitions préalables	167
2 L'éveil de la discipline au Canada	169
3 Les premiers arrangements institutionnels d'évaluation (1977) : la mise en place "formelle" de l'évaluation au gouvernement fédéral	171
4 La création du Bureau du contrôleur général du Canada	171
5 La situation au début des années 1990	172
6 1994 : de nouvelles orientations en matière d'évaluation	173
7 La politique d'examen	174
8 L'évaluation hors des ministères fédéraux et du Conseil du Trésor du Canada	178
9 Conclusion	179
9.1 Les gestionnaires de programme.....	180
9.2 Les parlementaires.....	180
9.3 Le citoyen canadien.....	181
10 Références bibliographiques	183

Contribution VII Guide des meilleures pratiques à suivre pour l'évaluation par l'OCDE..... 189

• RÉALISER DES EVALUATIONS EFFICACES.....	190
1 Optimiser les évaluations	190
1.1 Définition et objectifs	190

1.2 Identifier les principaux participants	191
1.3 Formuler les avantages et les coûts	191
2 Organiser le cadre de l'évaluation	191
2.1 Favoriser la culture de l'évaluation	191
2.2 Gestion stratégique des activités d'évaluation	192
2.3 Améliorer la crédibilité.....	193
3 Réaliser des évaluations efficaces.....	193
3.1 Assurer des liens avec les processus de prise de décision.....	193
3.2 Choisir les évaluations qui conviennent	194
4 Au sujet de cette note de synthèse... ..	195

<u>Liste des membres du Conseil national de l'évaluation (décret du 26 janvier 1999).....</u>	<u>197</u>
<i>Membre du Conseil d'État :</i>	198
<i>Membre de la Cour des comptes :</i>	198
<i>Membres du Conseil économique et social :</i>	198
<i>Maire :</i>	198
<i>Conseiller général :</i>	198
<i>Conseiller régional :</i>	198

Introduction

Le rapport du Conseil national de l'évaluation pour l'année 1999 comporte deux parties.

Première partie : le rapport d'activité proprement dit comprenant :

- Chapitre Premier : une analyse de l'activité du Conseil.
- Chapitre II : un compte rendu des conclusions méthodologiques que le CNE a tirées de l'examen des premiers projets retenus :
 - le projet d'évaluation concernant la lutte contre le Sida ;
 - le projet d'évaluation du programme « nouveaux services – nouveaux emplois dans le secteur de la Jeunesse et des Sports » et le projet d'évaluation des emplois non marchands ;

- le projet d'évaluation de la politique de préservation de la ressource destinée à la production d'eau destinée à la consommation humaine ;
- le projet d'évaluation des effets de la PAC sur l'environnement ;
- le projet d'évaluation de la politique interministérielle en faveur du logement social dans les départements d'Outre-mer. Ce chapitre explicite aussi les pistes à développer pour l'avenir.
- Annexe I : le discours du ministre de la Fonction publique, de la Réforme de l'État et de la Décentralisation lors de l'installation du CNE.
- Annexe II : la lettre du CNE transmettant au Premier ministre les projets d'évaluation retenus par le CNE et l'analyse de ces projets.
- Annexe III : la liste des instances en cours, présidents et rapporteurs.

Deuxième partie : un état de la réflexion sur la méthodologie et les pratiques de l'évaluation à travers différentes contributions nationales et internationales. Ces communications – qui n'engagent que leurs auteurs – ont été sélectionnées pour leurs contenus.

- Contribution I : les contributions d'experts français qui sont intervenus lors d'un séminaire méthodologique en novembre 1999 (Eric Monnier et Jacques Toulemonde).
- Contribution II : Maurice Baslé analyse les pratiques de l'évaluation française des contrats de Plan État-région.
- Contribution III : Sylvie Trosa résume les tendances internationales en matière de développement de l'évaluation.
- Contribution IV : Christopher Pollitt retrace un débat important au sein de la Société européenne d'évaluation sur l'équilibre à trouver entre évaluation participative et évaluation scientifique. Le CNE qui en a débattu a conclu que ces deux formes d'évaluation demeurent largement nécessaires et complémentaires.
- Contribution V : Keith Mackay qui, en exposant le développement de l'évaluation en Australie, explique l'intérêt d'une liaison entre l'évaluation et les décisions budgétaires.
- Contribution VI : Jean Turgeon présente l'évaluation des politiques et des programmes au niveau du gouvernement fédéral canadien.
- Contribution VII : le guide de l'OCDE donne des conseils de bonne pratique en matière d'évaluation.

L'ensemble de ces contributions ne prétend pas résumer les débats actuels mais rendre compte des principales tendances où s'inscrit la réflexion du champ français de l'évaluation.

Chapitre Premier

La nouvelle procédure

Le décret du 18 novembre 1998 instituant le Conseil national de l'évaluation et modifiant le dispositif antérieur du Conseil scientifique de l'évaluation, issu du décret du 22 janvier 1990, a introduit une nouvelle procédure. Les projets d'évaluation interministériels de politiques publiques, proposés par l'administration, sont analysés par le Conseil et remis au Premier ministre, qui arrête le programme annuel des évaluations.

1. L'activité du Conseil en 1999

L'année de la mise en place du Conseil national de l'évaluation, on dénote deux tendances opposées.

D'une part, le nombre élevé de propositions d'évaluation présentées par les administrations marque un intérêt réel pour la matière. Toutefois on note la sur-représentation des projets venant d'administrations plus « faibles ».

À l'inverse, pour certains hauts fonctionnaires l'évaluation reste perçue comme un exercice de recherche ésotérique et long, alors que le nombre de domaines où l'administration manque d'évaluations, c'est-à-dire de données fiables, de faits et de connaissance du terrain est élevé. Il y a donc une tendance à se plaindre de la complexité de l'évaluation tout en admettant que la connaissance des réalités de la mise en œuvre des politiques publiques est trop parcellaire. Par exemple, qui sait pourquoi les dispositifs juridiques en matière d'eau potable et de protection de la ressource ne sont pas toujours appliqués ? Qui connaît les alternatives aux mécanismes actuels de prévention et de traitement du SIDA ? Qui a évalué objectivement les nouvelles formes d'aides à l'emploi ?

Retenons-en un constat optimiste : il existe un intérêt réel dans les administrations pour l'évaluation. Mais mesurons aussi l'ampleur de l'effort

encore à effectuer pour convaincre les fonctionnaires de tous niveaux que l'évaluation peut leur être utile.

Nous faisons encore trop confiance à l'intelligence individuelle au détriment de la recherche patiente et collective de faits appuyés sur des données fiables. Nous avons été éduqués dans un système qui valorise les compétences individuelles, face à l'esprit d'équipe, et l'idée brillante face au raisonnement basé sur des démonstrations empiriquement vérifiables. Un peu plus d'intelligence collective, un peu plus de recherche empirique, tels sont les souhaits du CNE.

Le décret du 18 novembre 1998 définit l'évaluation comme ayant « pour objet d'apprécier, dans un cadre interministériel, l'efficacité de cette politique en comparant ses résultats aux objectifs assignés et aux moyens mis en œuvre ».

Le Conseil national est composé de quatorze membres nommés pour trois ans dans les conditions suivantes : six personnalités qualifiées choisies en raison de leurs compétences en matière d'évaluation ; un membre du Conseil d'État et un membre de la Cour des comptes désignés par ces institutions ; trois membres du Conseil économique et social désignés par celui-ci ; un maire, un conseiller général et un conseiller régional au vu des propositions faites par leurs organisations représentatives.

Le président du Conseil est nommé par décret, parmi ses membres, sur proposition du Conseil. Il est assisté d'un rapporteur général, nommé par arrêté du Premier ministre, sur proposition du Commissaire au Plan.

Le Conseil a été mis en place par le ministre de la Fonction publique, de la Réforme de l'État et de la Décentralisation le 16 février 1999.

Depuis il s'est réuni huit fois, le principe étant une réunion mensuelle.

Le programme d'évaluation établi pour l'année 1999 a été transmis au Premier ministre le 30 juin 1999 et comprenait six projets d'évaluation, dont cinq ont été retenus. À ce jour, toutes les instances d'évaluation ont été mises en place et ont commencé leurs travaux.

Le CNE s'est également réuni en séminaire le 25 novembre pour définir sa propre doctrine quant à la méthodologie de conduite de l'évaluation et a auditionné différents experts retenus pour leur expérience internationale : Eric Monnier (C3E), Michel Setbon (CNRS) et Jacques Toulemonde (C3E).

Le *Petit guide de l'évaluation* élaboré par le précédent Conseil scientifique de l'évaluation a été adapté à la nouvelle situation du CNE et traduit en anglais et va être mis sur Internet pour le public non francophone sous le titre de *A Practical Guide to Program and Policy Evaluation*.

2. L'originalité du nouveau dispositif

Par rapport au précédent Conseil scientifique de l'évaluation, la composition du Conseil national a été modifiée : elle comprend des représentants des collectivités territoriales et des milieux socioprofessionnels représentés au Conseil économique et social.

Ses compétences ont également été élargies : le Conseil national peut être consulté non seulement par les administrations de l'État mais également par les collectivités territoriales et leurs établissements publics sur toute question méthodologique relative à la conduite d'une évaluation.

2.1. Une procédure simplifiée

Le dispositif est caractérisé par une simplification de la procédure comparée à la précédente. Auparavant les projets devaient être soumis à un comité interministériel ad hoc de ministres chargé de les examiner. Ce comité a été supprimé. À la place, le Conseil propose au Premier ministre un programme annuel qui présente les projets d'évaluation retenus, expose leur contenu ainsi que les raisons justifiant leur sélection, et précise les conditions dans lesquelles seront mises en œuvre les évaluations : le mode de composition de l'instance chargée de conduire l'évaluation, les modalités de sa mise en œuvre, les critères en vertu desquels seront choisis les opérateurs publics ou privés chargés de procéder à l'évaluation, le délai de sa réalisation, son coût et ses modalités de financement.

Au regard du précédent dispositif cette procédure comporte moins de phases : elle groupe les évaluations dans un programme annuel arrêté par le Premier ministre. Les administrations à l'origine des projets ne sont pas pour autant dessaisies du pouvoir de donner leur avis dans la mesure où le Commissariat général du Plan assure la coordination et la concertation interministérielle avant que le programme ne soit transmis au Premier ministre. Au cours de cette phase,

le Commissariat général du Plan cherche à privilégier l'accord des administrations autour du projet d'évaluation.

De fait les projets d'évaluation soumis au Premier ministre en juin 1999 ont résulté d'un processus de mûrissement et d'interactions positives avec les administrations. Tant le Commissariat général du Plan que le CNE y ont participé. Ce processus qui commence parfois avec de simples propositions peut déboucher, par itérations successives avec les administrations, en véritables projets d'évaluation à la méthodologie et aux problématiques claires.

2.2. Une étroite collaboration avec le Commissariat général du Plan

Le Commissariat général du Plan assure le secrétariat du CNE. Il prend en charge la coordination administrative, il instruit les projets d'évaluation avant la saisine du CNE, puis, une fois leur programme fixé, met en place les instances d'évaluation, suit les travaux d'évaluation et assure leur publication. Il a aussi mission de proposer au Premier ministre les suites à donner aux évaluations pour ce qui concerne l'État. Il favorise le développement de l'évaluation, notamment en rassemblant et en diffusant l'information relative aux méthodes et techniques d'évaluation pratiquées en France et à l'étranger. Il contribue à la formation dans le domaine de l'évaluation. Il gère le Fonds national de développement de l'évaluation.

2.3. Le maintien du contrôle de la qualité des évaluations

Point commun avec l'ancien CSE, le CNE a la mission de vérifier la qualité des évaluations, et apprécier notamment la compatibilité de leurs recommandations avec les travaux et études effectués. Les instances d'évaluation transmettent leurs rapports au Conseil national qui dispose de deux mois pour émettre son avis. C'est ensuite le CNE qui adresse les rapports d'évaluation ainsi que son avis aux administrations, collectivités ou établissements concernés qui lui font connaître dans un délai de trois mois les suites qu'ils envisagent de donner à ces rapports. Au terme de ce délai les rapports d'évaluation sont rendus publics et comportent en annexe les éléments du programme d'évaluation, l'avis du CNE et les réponses des administrations, collectivités ou établissements publics concernés.

2.4. Les étapes de la pratique actuelle en 1999

Cette procédure a suivi en 1999 les étapes suivantes :

Phase un

Faisant suite à la circulaire du Premier ministre du 7 septembre 1998, le Commissariat général du Plan a sollicité les administrations pour qu'elles transmettent des propositions d'évaluation. Il a aidé les administrations à préciser leurs projets afin que ceux-ci répondent aux critères que l'on trouve communément admis tels qu'ils sont décrits notamment dans le *Petit guide de l'évaluation* élaboré par l'ancien CSE : clarté des objectifs de la politique publique et de l'évaluation, des problèmes à évaluer, de la méthode d'investigation envisagée.

Phase deux

Le Commissariat général du Plan saisit ensuite le CNE des propositions d'évaluation. Le CNE a également la capacité de proposer des thèmes ou des axes prioritaires d'évaluation. Cette capacité n'a pas été employée en 1999.

Une fois saisi, le CNE désigne des rapporteurs en son sein qui, en liaison avec les administrations, donnent des conseils sur le champ et la pertinence des projets, les méthodes proposées et aident les projets d'évaluation dans leur élaboration.

Le CNE a privilégié une approche collégiale. Il a désigné en son sein, pour chaque projet, un « binôme » de rapporteurs. Ceux-ci rendent compte au Conseil qui décide ou non de mettre le projet d'évaluation concerné à l'agenda du programme d'évaluation transmis au Premier ministre.

Le CNE a transmis sa proposition de programme ¹ le 29 juin 1999 au Premier ministre. Celui-ci a décidé cinq évaluations en juillet.

Le tableau ci-après présente le processus de préparation des projets d'évaluation, de la circulaire du Premier ministre en date du 7 septembre 1998 appelant les ministères à faire connaître au Commissaire général du Plan leurs intentions d'évaluation à la décision du Premier ministre lors du CIRE du 13 juillet 1999. Partant de 37 intentions d'évaluation, un premier examen a

(1) Cf. lettre de transmission et ses annexes, en annexe II.

permis de retenir quatorze dossiers ; parmi ceux-ci, six ont été menés à bien, deux n'ont pu donner lieu à l'élaboration d'un cahier des charges satisfaisant. Neuf intentions retenues par le CNE sont actuellement travaillées pour devenir des projets d'évaluation. Dès lors que les projets seront arrivés à maturité et auront été discutés en réunions inter-administratives, le CNE formalisera son nouveau programme à l'intention du Premier ministre.

Phase trois

Les cinq instances d'évaluation ont été préparées puis installées par le Commissariat général du Plan entre novembre 1999 et février 2000, à l'exception d'une instance qui a été mise en place en avril 2000.

Une audition est prévue par le CNE du président et des rapporteurs des instances après le démarrage de leurs travaux. Le CNE considère cette phase comme utile et constructive car elle ouvre un dialogue d'ordre méthodologique avec les responsables des instances, leur permet d'énoncer leurs propres préoccupations et facilite la prise en compte des suggestions du CNE.

La durée des travaux n'excèdera pas 18 mois.

Phase quatre

Une fois les travaux terminés, les instances transmettront leur rapport au CNE, qui rendra un avis sur leur qualité dans un délai de deux mois.

Cet avis sera transmis aux administrations ou collectivités concernées, qui auront trois mois pour répondre et préciser les suites envisagées.

L'ensemble des documents afférents à l'évaluation, rapport, études, avis du CNE et réponses des administrations et des collectivités sera ensuite rendu public. Cette publication sera assurée par le Commissariat général du Plan.

Phase cinq

Le Commissaire au Plan proposera au Premier ministre, en étroite liaison avec les administrations, les suites à donner aux évaluations en ce qui concerne l'État.

Chapitre II

Les conclusions méthodologiques tirées des premiers projets

Ce chapitre résume les conclusions et conseils méthodologiques que le CNE a émis lors de l'examen des premiers projets d'évaluation qui lui ont été soumis.

Le CNE a d'abord été favorablement impressionné par la qualité des projets d'évaluation qui lui ont été soumis. Toutefois, un certain nombre de remarques ont été faites sur quasiment l'ensemble des projets. Celles-ci, qui seront développées dans la suite de ce rapport, résument les conseils que le CNE a donnés aux administrations. Elles se fondent sur le principe que l'évaluation n'est pas un processus d'étude et de recherche mais est insérée dans un cadre de décision politique (d'où l'importance de comprendre les motivations du commanditaire) ou de changement social (d'où l'importance de l'écoute et de l'association des parties prenantes).

1. Des conclusions stratégiques

L'évaluation recouvre différentes pratiques qui vont de l'évaluation courte et ciblée aux investigations de causalité complexes.

Le CNE s'est inscrit dans la philosophie de la circulaire du Premier ministre du 28 décembre 1998 dont l'objectif est de privilégier les évaluations courtes, de préférence d'un an, d'au plus 18 mois. En effet, certaines évaluations de la période précédente se sont étendues sur plus de cinq ans ce qui leur a enlevé une bonne part de leur capacité d'influencer les décisions politiques. Lorsqu'une évaluation dure plus d'un an, le CNE insiste pour que des points d'étape réguliers lui parviennent et que ceux-ci soient adressés aux commanditaires de l'évaluation.

Néanmoins les évaluations résultent d'un jeu sur des contraintes objectives qui sont temporelles et psychologiques.

1.1. Le temps nécessaire à une évaluation

Une évaluation nécessite un certain nombre de phases de réflexion en amont.

- Le temps d'élaborer des projets viables lorsqu'ils émanent d'administrations qui n'ont pas toujours de compétences spécifiques en matière d'évaluation.
- Le temps de la mise en place des instances d'évaluation qui font appel à des participants bénévoles. Ceux-ci doivent à la fois connaître le sujet en cause et n'être pas directement impliqués dans le jeu d'intérêts entourant la politique à évaluer, ce qui n'est pas une équation simple à mettre en œuvre car il arrive que ceux qui ont une bonne connaissance technique soient aussi attachés au maintien des positions de leur administration ou groupe d'origine.
- Le temps des travaux qui, lorsqu'ils touchent des évaluations mettant en cause différents niveaux de collectivités publiques (Union européenne, région, département et/ou commune), peuvent nécessiter la mise en place de groupes de travail locaux et la réalisation de travaux de terrain qui requièrent le temps de la concertation et de l'exécution. Le temps à y consacrer peut aussi avoir un effet positif, celui de la maturation des prises de conscience ou d'une nouvelle culture de la politique publique.
- De plus, si certaines évaluations dans le passé ont été longues, leur sérieux sert aujourd'hui de point de référence aux réflexions et actions actuelles sur une politique publique. Il en a été ainsi de l'évaluation sur la prévention des risques naturels.
- Les évaluations posent la question de la mémoire collective de l'administration. Celle-ci a tendance à se dissiper trop vite soit parce que les personnels changent de poste soit parce que les cloisonnements internes font que l'information circule mal non seulement d'une administration à l'autre mais au sein d'une même administration. Le « marketing » des évaluations devient alors un enjeu ; il serait facilité par l'existence d'un correspondant évaluation ou d'un bureau de l'évaluation dans chaque ministère.

Parfois la durée traduit, plus fondamentalement, l'usage fait des évaluations. Ce peut être l'apprentissage collectif des acteurs de l'évaluation qui, en cours même d'évaluation, apprennent à changer leurs pratiques ou leurs relations de

travail. Mais cet usage, pour utile qu'il soit, doit être suivi de décisions politiques. Chaque évaluation doit gérer ces aspects stratégiques dans son pilotage.

La durée optimale d'une évaluation résulte du meilleur compromis entre :

- les attentes du commanditaire politique et le temps nécessaire à un travail rigoureux et scientifique ;
- l'apprentissage collectif et la recherche de données ;
- l'accent mis sur la décision « post évaluation » et le caractère formateur de l'évaluation au cours du processus de débat et de concertation qu'elle enclenche.

Chaque évaluation appelle un arbitrage dès son début. Qu'est-ce qui est le plus important ? Répondre rapidement au commanditaire car le sujet est urgent ou au contraire capitaliser le plus de savoir scientifique pour établir un travail qui servira de cadre de référence sur le moyen et long terme ? Ou encore convaincre les acteurs qu'il faut changer leurs façons de faire ?

La durée, symptôme d'un problème plus large, dénote la grande diversité des évaluations selon les sujets qu'elles abordent. Savoir si des policiers armés sont plus efficaces qu'une force de police sans armes suppose des analyses de causalité complexes, des échantillons comparant les différentes forces de police, des instruments de mesure sophistiqués. En revanche, évaluer l'impact d'une action de formation ne peut requérir que l'élaboration et la mesure de quelques indicateurs et un groupe de travail les interprétant. Dans les deux cas, la durée ne sera pas la même. Donc, se préoccuper de la durée revient à faire passer le message que *toutes les évaluations ne sont pas lourdes et complexes et qu'il est possible d'évaluer au quotidien.*

L'idéal serait de marier dans le programme du CNE des évaluations « courtes et ciblées », des évaluations à visée scientifique et des évaluations qui permettent de rapprocher les cultures des parties prenantes, et donc de jouer sur différents niveaux de complexité et de durée.

1.2. Des évaluations au service de l'avenir

L'évaluation demeure encore perçue dans certaines administrations comme :

- . la recherche du blâme et de la faute ;
- . une analyse purement rétrospective ;
- . une mise en lumière exclusive des dysfonctionnements.

Un signe clair de cette culture est donné par certains des premiers projets transmis au CNE, focalisés sur les points noirs de l'action administrative, alors qu'il peut être tout aussi intéressant (mais peut-être plus complexe) d'évaluer les réussites et pourquoi ce sont des réussites.

À notre sens, les évaluations doivent se situer dans un esprit qui n'est pas celui de l'incrimination mais de l'explication des causes des résultats observés, des effets de système ou de structure. S'il est important de comprendre les dysfonctionnements, il est tout aussi important de mettre en valeur certaines pratiques efficaces de l'action administrative.

Enfin l'évaluation, analyse rétrospective, doit être au service de l'avenir. Elle doit pouvoir aider les administrations à mieux gérer, ou rapprocher dans leur pratique les acteurs administratifs, socio-professionnels et les usagers. En toute hypothèse, il est souhaitable que l'évaluation pose les jalons d'une meilleure action publique pour le futur, ce qu'elle peut faire par :

- la mise en évidence d'indicateurs de pilotage ;
- la conception de systèmes d'information ;
- la formalisation de leçons qui peuvent être transposées à d'autres politiques publiques du même type.

1.3. L'évaluation, partie intégrante d'une culture du résultat

La seconde conséquence du lien nécessaire entre évaluation et modernisation du service public est d'intégrer l'évaluation dès le commencement des processus de modernisation et de ne pas en faire une discipline ou une activité à part. Contrôle de gestion, démarche qualité, « benchmarking »¹, évaluation se complètent.

(1) Le « benchmarking » ou parangonage – selon le vocable de l'Académie - recouvre des activités diverses qui consistent à rechercher des exemples de « bonne pratique » afin d'améliorer les siennes propres (comparaison pour l'amélioration de l'efficacité), à des procédures d'étalonnage où les performances de services comparables sont utilisées pour classer ces services en bons ou mauvais.

Comparée aux pays anglo-saxons, la France a eu trop tendance, peut-être sur le modèle des évaluations lourdes du General Accounting Office aux États-Unis, à ne pas promouvoir une *culture de la performance* dont l'évaluation est en quelque sorte un réflexe immédiat : lorsque l'administration se fixe des objectifs, elle devrait mettre en place aussitôt un plan de mise en œuvre et un système d'information et commencer à évaluer dès le début de l'action publique nouvelle décidée.

Cette évaluation concomitante à la mise en œuvre de l'action publique peut être l'affaire d'un comité de pilotage de suivi. Elle peut aussi être plus élaborée avec un contenu méthodologique et l'appel à des experts externes.

Cette culture de la performance doit s'appliquer à toutes les actions publiques. Elle permet de banaliser l'évaluation comme une procédure nécessaire, courante et faisant partie de la gestion publique. Ceci, bien entendu, ne remet pas en cause la nécessité d'évaluations complexes recherchant les causes des résultats analysés d'une politique. Toutefois, ces évaluations plus complexes seront d'autant mieux acceptées qu'elles feront partie d'une culture généralisée de l'évaluation et de la modernisation.

Ainsi, la loi qui avait institué le RMI avait prévu son évaluation, le budget de celle-ci et le profil de l'instance qui en serait chargée. C'est un exemple de bonne programmation de l'action publique.

1.4. L'évaluation est plus qu'un bilan ou un constat, c'est la compréhension des causes des résultats

Que l'évaluation soit « simple » (évaluation d'un plan de formation) ou complexe (efficacité des aides pour l'emploi) le CNE estime qu'il n'y a pas évaluation tant que la question de la causalité (c'est-à-dire pourquoi avons-nous obtenu les résultats que nous avons mesurés et analysés ?) n'est pas posée. L'évaluation n'est pas une liste de résultats mais une interrogation sur le « pourquoi » : pourquoi tel résultat existe-t-il ? S'agit-il de problèmes de mise en œuvre, de stratégie des acteurs, d'une incapacité d'appliquer des objectifs ? Les causes peuvent être très différentes, mais l'évaluation commence avec l'analyse objective des causes.

Le séminaire du 25 novembre tenu par le CNE a permis d'établir un constat, certes non exhaustif, des pratiques d'évaluation au niveau européen. Trop d'évaluations ne sont encore en fait que des bilans alors qu'elles doivent être des analyses de causalité établissant les liens entre les résultats mesurés et les objectifs initiaux.

À cet effet le CNE s'est penché sur l'exemple de la loi sur la montagne. Le rapporteur de l'ex-CSE a donné sa perception de la façon dont s'est déroulée l'évaluation de la loi montagne. Décidée le 17 mars 1993, elle a été transmise au CSE le 23 décembre 1993 et le CSE a rendu son premier avis le 14 avril 1994. La première réunion de l'instance a eu lieu le 6 avril 1995 avec un premier rapport d'étape au CSE en juin 1996, une première version du rapport en juin 1998 et une publication en novembre 1999. Soit six ans, dont deux ans et demi de travail effectif. Cette durée est trop longue. L'explication est à chercher dans la complexité des arbitrages pour constituer l'instance.

D'après le rapporteur de l'ex-CSE, ce rapport donne un bilan assez complet de la situation économique et sociale de la montagne et des évolutions historiques qui y ont conduit, mais ne donne que peu d'informations sur les liens de causalité entre cette situation et les actions publiques menées. L'évaluation du dispositif institutionnel demeure parcellaire. En fait, même si l'évaluation conduite amène à discerner de bons résultats de l'action publique (la biodiversité a été préservée, les subventions sont moindres que dans d'autres zones en difficulté), elle préconise le renforcement du dispositif montagne, essentiellement parce que l'instance n'était composée que de membres ayant intérêt à la consolidation de ce dispositif. D'ailleurs, une des recommandations du CSE, qui était de faire des monographies de terrain pour comprendre comment cela se passait in situ n'a jamais été suivie. Il n'y a pas eu non plus de bilan financier.

Il aurait fallu sélectionner quelques mécanismes d'action particuliers et approfondir leur analyse, quitte à reconnaître que l'ensemble du champ n'était pas couvert.

C'est une des principales difficultés des instances d'évaluation. Parce qu'elles veulent bien faire, elles embrassent trop large et n'arrivent pas à traiter l'ensemble des questions qui se posent à elle. L'expérience montre qu'il vaut mieux partir de quelques hypothèses ciblées sur des mécanismes d'action, les analyser, prouver et vérifier, quitte à n'élargir l'interrogation que dans un second temps.

L'un des rapporteurs de l'instance explique les difficultés qui ont jalonné cette évaluation :

- . la complexité des évaluations territoriales parce que l'on couvre un champ tellement vaste que l'instance perd de vue les objectifs premiers de l'évaluation (ce qui justifie la nécessité de limiter le nombre des interrogations comme le recommande le CNE) ;

- . dans l'instance, peu de personnes savaient vraiment comment conduire une évaluation et la plupart confondait cet exercice avec celui d'un bilan ;

- . aucune grille d'écriture et d'analyse commune n'a été donnée aux différents rapporteurs ;

- . l'instance n'a pas construit d'indicateurs de réussite ou d'échec préalable ;

- . l'analyse financière a été impossible parce que les chiffres étaient disséminés dans de multiples endroits sans être identifiés à des services rendus.

Dès lors, quelles conclusions retenir de cette évaluation de la politique de la montagne ?

- . la nécessité de bien définir le champ de l'évaluation ;

- . l'importance de former les membres des instances aux concepts de base de l'évaluation ;

- . la constitution d'un référent commun ;

- . la collaboration des services financiers pour avoir au moins une certaine idée des montants concernés.

1.5. L'évaluation nécessite des instances d'évaluation pluralistes, compétentes et efficaces

Le CNE a consacré une partie de ses travaux aux questions du rôle et de la constitution des instances d'évaluation. Celles-ci remplissent un double rôle qui est d'assurer la constitution d'une culture commune aux membres de l'instance, notamment aux administrations qui la composent et qui n'ont pas toujours l'occasion de travailler ensemble, dans un contexte qui n'est pas celui de la décision, mais aussi d'assurer un pilotage efficace des travaux, ces deux buts n'étant pas toujours faciles à concilier. En effet, l'efficacité est plus facilement

soutenue par un comité de pilotage léger et uni alors que l'apprentissage collectif des administrations requiert des instances plus larges dont une partie du travail sera moins le pilotage des travaux que la mise en harmonie des acteurs administratifs.

La composition des instances est un exercice difficile en raison de la charge de travail qu'implique la présidence et même la participation aux instances. La désignation des présidents remonte (sans nécessité) au plus haut niveau de chaque ministère, ce qui entraîne un délai supplémentaire. Le Conseil national de l'évaluation a rendu hommage au Commissariat général du Plan pour la célérité de ce travail de composition des instances, compte tenu des contraintes ; en 1999, si l'on exclut la période des vacances, le délai moyen a en effet été de trois mois.

Le CNE a voulu s'appuyer sur l'expérience du CSE pour tirer ses propres conclusions quant aux instances d'évaluation. Cette expérience a montré les raisons du dysfonctionnement de certaines instances de la période CSE.

- Elles ont, d'abord, compris trop de membres. Dès lors les accords à tisser entre les participants ont pris beaucoup de temps, plus que le pilotage et l'analyse des travaux effectués.
- Elles étaient au surplus composées tant du côté des administrations que des experts, de convaincus voire de militants de la politique publique en cause ; dès lors, elles se sont comportées plus comme des porte-parole et des défenseurs que comme des analystes ayant une distance critique et objective face à la politique à évaluer.

Lors du séminaire du 25 novembre, Michel Setbon (CNRS) qui a réalisé avec Pierre Lascoumes (CNRS), à la demande de l'ancien CSE, une analyse critique de six évaluations en cours ou passées a d'abord souligné que leur recherche ne portait pas sur tous les types d'évaluation, mais sur les cas les plus complexes c'est-à-dire les évaluations de politiques publiques interministérielles : ces évaluations sont plus complexes parce qu'elles impliquent de nombreux acteurs et ont un effet direct sur le politique.

L'interrogation a porté sur les effets du choix d'une instance pluraliste, en notant d'ailleurs que dans la plupart des cas, les membres étaient des représentants de l'administration (pluralisme interne à l'administration) et que de plus la composition de l'instance pouvait varier d'une réunion à l'autre, les fonctionnaires se faisant représenter ou les présidents étant changés et nommés à d'autres fonctions. Il n'y a donc pas eu de constitution d'équipe stable.

Les conclusions de cette recherche ont été « en temps réel » :

- une absence de définition de l'objet d'évaluation faute d'accord sur les objectifs de l'évaluation entre des partenaires de l'instance qui se sont vécus comme représentants « de leur ministère » ;
- une méfiance généralisée par rapport à l'évaluation, avec la crainte qu'elle ne soit utilisée contre ses commanditaires ;
- une ambivalence fondamentale dans la mesure où dans bien des cas le commanditaire s'est aussi trouvé être l'acteur principal de l'instance ;
- le CSE a été un acteur ignoré dans la mesure où ses recommandations n'ont pas été comprises par les instances faute de l'instauration d'un dialogue constructif.

Dans la mesure où la présence de scientifiques dans les instances a parfois été très marginale (« un » seul scientifique ne peut ni contrebalancer le poids des administratifs ni assurer la « traduction » des études parfois complexes en recommandations administratives), nombre d'instances n'ont pu intégrer les données d'équipes d'évaluation venant d'expertises très différentes.

Certains rapports n'ont en fait que rassemblé des recommandations suggérées dès le départ par le commanditaire, et y ont ajouté quelques concessions portant sur des dispositifs à améliorer.

Les évaluations les plus dynamiques ont probablement été celles où des processus d'apprentissage locaux entre acteurs publics ont modifié les pratiques locales.

Ainsi, par exemple, l'évaluation des programmes régionaux de formation professionnelle a été pilotée par un comité tripartite avec les administrations de l'État, les partenaires sociaux et les régions. Cette évaluation, grâce au dialogue établi entre les principaux partenaires, a permis de mettre en lumière des conclusions importantes pour l'amélioration de cette politique.

En résumé, le CNE préconise des instances d'évaluation :

- . qui comportent moins de membres afin d'assurer l'efficacité du pilotage ;
- . qui ne soient pas dominées par les parties prenantes qui ont un intérêt direct dans la défense des positions de leur administration ou groupe d'origine ;
- . dont les membres soient à la fois crédibles, reconnus et compétents.

Par ailleurs, la nécessité qui est souvent avancée d'associer toutes les administrations concernées, faute de quoi elles ne fourniraient pas d'informations pertinentes et bloqueraient l'accès aux sources, ne laisse pas de surprendre, dans la mesure où toute évaluation est une décision du Premier ministre et où l'hypothèse implicite serait que les administrations ne joueraient pas le jeu. L'expérience montre que les administrations ne laissent pas accès à leurs données que si elles sont étroitement associées à la conduite de l'évaluation. Le danger n'en demeure pas moins de ne voir se développer que des auto-évaluations, justifiant la situation existante ou ne l'améliorant qu'à la marge.

D'autres moyens d'avoir un débat large et pluriel que la présence de toutes les administrations concernées, peuvent être imaginés : sous-groupes de travail, auditions, réunions locales, etc.

In fine, il apparaît nécessaire qu'un « noyau dur homogène » constitué d'un petit groupe de personnes assure le pilotage au quotidien tout en se donnant les moyens d'entendre les points de vue d'un éventail large de participants et d'usagers ayant une opinion sur la politique publique à évaluer.

Il est à noter que, pour les cinq premières évaluations, le Commissariat général du Plan a déployé des efforts importants pour limiter le nombre des membres des instances d'évaluation et s'assurer de la présence d'experts indépendants.

De même que le CSE, le CNE insiste sur les exigences d'une évaluation indépendante. Une évaluation doit pouvoir poser les questions fondamentales, voire difficiles. Pour ce faire, il est important que l'instance d'évaluation ne subisse pas de pressions dans le cours de son travail et que ses membres agissent *intuitu personae* et non comme des mandataires de leur organisation. L'instance d'évaluation est d'abord un lieu de travail et d'acculturation plutôt qu'un lieu de négociation et d'arbitrage.

2. Des conclusions méthodologiques

Il faut mieux définir les objectifs de l'évaluation et mettre en lumière l'importance d'un diagnostic préalable et l'existence d'hypothèses expliquant les dysfonctionnements.

2.1. Où est le problème?

Nombre de projets ont bien décrit la politique publique à évaluer mais de manière descriptive sans expliquer « quel est le problème » c'est-à-dire pourquoi le projet d'évaluation est amorcé. Sans la compréhension de ce qui fait problème, il est difficile de connaître les raisons du commanditaire pour engager l'évaluation. Pourquoi évaluer s'il n'y a pas désir de comprendre ou de progresser ? Le danger de projets trop ouverts, non guidés par la formulation du problème, est d'apporter des réponses qui ne satisferont pas les attentes du commanditaire ou de s'enliser dans des travaux trop ambitieux.

Certains projets se sont présentés comme des listes de questions plutôt que comme des hypothèses à vérifier. Si chaque évaluation part d'un constat (à vérifier) de dysfonctionnements, il est nécessaire de connaître les hypothèses du commanditaire quant aux raisons de ces dysfonctionnements. Les hypothèses se différencient d'une liste de questions car elles formulent des causes possibles des dysfonctionnements et vont servir à l'analyse. Encore une fois, l'évaluation n'est pas de la recherche pure mais un travail d'analyse objective qui s'insère dans le processus de décision politique et administratif. Des hypothèses pourront être infirmées, confirmées ou retravaillées. Elles établissent un dialogue avec le commanditaire.

En revanche, si le projet d'évaluation n'est qu'une liste de questions, toutes les réponses sont possibles et se valent. Des hypothèses, au contraire, servent de base à un travail cadré et rigoureux car elles fournissent un point de départ aux investigations et permettent à l'évaluation, si besoin est, de faire évoluer les mentalités.

Éviter les jugements arbitraires

Sans formalisation de ce qui est considéré comme succès ou comme échec par les commanditaires (c'est-à-dire en oubliant la mise en place d'indicateurs) les jugements portés par les travaux d'évaluation risquent d'être arbitraires et non référés à une norme de jugement claire et connue. Ce point peut n'être que partiellement résolu au stade du projet d'évaluation car la construction d'indicateurs est bien souvent le premier travail qui permet de souder l'instance d'évaluation. Que ces indicateurs préexistent à la mise en place de l'instance d'évaluation ou qu'ils soient élaborés par cette dernière, ce qui compte c'est que les indicateurs soient construits.

En fait, toute évaluation doit partir de critères ou de normes de références sauf à se contenter d'opinions subjectives. On ne peut qualifier une action publique de bonne ou de mauvaise que si l'on a été clair sur les critères qui amènent à la qualifier comme telle. Ces normes se traduisent souvent par des indicateurs qu'il faut eux-mêmes interpréter et cette interprétation est de nature normative. Par exemple, sur certaines aides pour l'emploi, il est apparu au vu de l'évaluation que l'effet d'aubaine a été important : près de 80 % des emplois auraient été créés sans l'aide. La question devient alors : accepte-t-on que 20 % seulement de nouveaux emplois aient été créés grâce au financement de 100 % ? En ce domaine, il n'existe pas de réponse juste ou fautive mais des réponses adéquates aux objectifs que le commanditaire cherche à poursuivre. L'évaluation qui a mis en évidence le fait que 100 % de subventions ne créaient que 20 % d'emplois nouveaux doit dresser ce constat, peut faire des recommandations pour améliorer la situation mais doit fonder son jugement sur les objectifs initiaux de la politique publique.

Mettre à plat les facteurs explicatifs

L'évaluation consiste à discerner les facteurs explicatifs des résultats des politiques publiques. Les indicateurs de succès ou d'échec de la politique publique aident à mettre à plat les facteurs explicatifs des résultats observés car ils définissent ce qui était attendu de la politique publique et ce qui n'en était pas attendu. La mise en évidence des facteurs explicatifs n'est pas aisée dans la mesure où il est parfois difficile de cerner ce qui est attribuable à l'action publique et ce qui peut être attribué à des effets externes (taux de croissance, taux de change, etc.). Néanmoins, la mise à plat de l'ensemble des facteurs explicatifs est déjà un progrès en soi et peut permettre de tirer quelques conclusions du type « dans tel contexte, l'action publique a produit tels effets ».

En conclusion, l'évaluation des politiques publiques est d'autant plus aisée qu'il existe déjà des systèmes d'information appuyés sur des indicateurs pertinents. L'évaluation ne saurait se résumer à la mesure mais est considérablement facilitée si elle dispose de données quantitatives et qualitatives mises en place en cours de processus de modernisation. Ceci n'est pas à ce jour réalisé.

Expliciter la logique d'intervention

Le CNE conseille aux administrations de penser les politiques publiques en termes de système c'est-à-dire de passer des objectifs généraux à des objectifs

concrets et spécifiques, mesurés à travers des systèmes d'information, alimentés par des indicateurs qui montrent si les objectifs sont réalisés, et complétés par une évaluation qui analyse les causes de l'écart entre les objectifs et les résultats. Cette forme de raisonnement, appelée communément dans les pays anglo-saxons *logique d'intervention des programmes* est la base nécessaire de l'évaluation et de la modernisation qui permet de mesurer les relations entre les objectifs, les moyens et les résultats. La logique d'intervention permet de mieux comprendre la politique publique avant de l'évaluer et de servir d'outil futur de pilotage de cette politique par les administrations en cause. Elle permet également de pallier l'insuffisante clarté des objectifs politiques de plus haut niveau, en définissant des objectifs pragmatiques et mesurables.

Objectifs généraux ou politiques ;
Objectifs spécifiques, mesurables et vérifiables ;
Stratégies de mise en œuvre de ces objectifs concrets ;
Mesures et système d'information ;
Indicateurs ;
Evaluation.

L'importance d'une mesure et d'une analyse des écarts entre les objectifs et les résultats

La logique d'intervention permet de mieux comprendre les objectifs qui ont guidé la politique publique. Le CNE a insisté sur l'importance du fait que les évaluations ne partent pas de questions ouvertes, mais d'hypothèses à confirmer ou infirmer sur le rapport entre les objectifs et les résultats, comme dans toute démarche scientifique. Nombre des évaluations qui n'ont pu aboutir par le passé, sont celles qui constamment étendaient leur champ à de nouveaux objets ou de nouvelles questions. Il en est ainsi de l'évaluation concernant l'aide aux enfants en difficulté qui, par-delà les institutions en cause, en est venue à s'interroger sur le rôle de l'école, puis sur celui de la famille, et n'a pu aboutir à des réponses ciblées.

Cela ne veut pas dire que l'évaluation doive se contenter de l'exploration des objectifs, mais qu'elle ne questionne leur pertinence qu'une fois que les raisons de l'écart entre les objectifs et les résultats ont été établies.

La vérification du couple objectifs/résultats est un moyen de canaliser l'évaluation dans un travail faisable en des délais limités et de la rendre utile au commanditaire, alors qu'une évaluation qui part de toutes les hypothèses

possibles (y compris celles qui n'ont pas été incluses dans les objectifs de la politique) risque d'aborder des questions impossibles à résoudre dans des délais raisonnables.

L'élargissement à de nouvelles sphères de questionnement qui ne sont pas incluses dans les objectifs, n'est nullement exclue mais apparaît comme une nouvelle étape face à l'analyse pragmatique de l'écart entre les objectifs et les résultats et aux raisons de cet écart.

Ces préconisations n'ont pas été conçues par le CNE comme un carcan visant à entraver la liberté de chercher et d'interroger mais comme un cadre permettant un travail utile. Il se peut que le diagnostic préalable, les hypothèses et les indicateurs soient remaniés par les travaux d'évaluation mais l'important est que le cadre de la politique publique dans lequel travaillent les évaluateurs ait été connu au départ.

2.2. Plus de quantitatif et plus d'écoute

Le rapport coût/efficacité

Les projets d'évaluation soumis au CNE en 1999 abordent rarement la question financière en raison des complexités de son investigation.

Il est vrai qu'en France il est difficile de se faire une idée quantitative d'une politique publique, c'est-à-dire de son coût budgétaire et analytique, car les cadres budgétaires et comptables ne correspondent pas à des services délivrés. De plus nombre de politiques publiques font l'objet de financement croisés.

Ce travail de connaissance des coûts s'avère toujours plus facile dans les pays anglo-saxons dont le budget fonctionne par programmes ou résultats, c'est-à-dire inclut une connaissance des coûts au moment du vote législatif des mesures et des politiques nouvelles, alors que le budget français demeure un budget par moyens et non de résultats.

Le CNE considère comme une lacune le fait de ne pas disposer, à la fin d'une évaluation, d'un bilan financier (quelles sommes pour quelle politique ?) même s'il ne s'agit que d'ordres de grandeur. Il peut en cela être aidé par les réformes en cours de la budgétisation et de la comptabilité publique qui permettent aux contrôleurs financiers de suivre en temps réel les droits constatés.

Pourtant il n'y a pas d'impossibilité : bien souvent un petit groupe de travail de haut niveau qui s'attacherait à définir quels sont les services délivrés (*outputs*)

et les impacts (*outcomes*) et qui travaillerait avec les services financiers au niveau territorial permettrait d'avoir une réelle approche des coûts ou des ressources utilisées.

Introduire l'international et le benchmarking (comparaison avec les pratiques les plus performantes)

Spontanément les premiers projets d'évaluation ne se sont pas posés la question de la comparaison avec des politiques publiques semblables menées dans d'autres pays. Or, pour pouvoir interpréter certains résultats quantitatifs et qualitatifs de l'évaluation, des procédures de comparaison internationale entre politiques semblables, ce que l'on appelle le « benchmarking » seraient utiles. Comment comprendre l'importance donnée par la France à la politique de la montagne si l'on n'a aucune donnée sur ce que font d'autres pays européens ? L'ouverture vers l'international permettrait à l'évaluation soit de découvrir d'autres façons de faire, soit de mieux justifier les choix effectués par la France.

L'évaluation est aussi l'écoute

La plupart des politiques publiques à évaluer mettent en jeu plusieurs niveaux de décision (Union européenne, État national, régions, départements, communes) et différents acteurs publics et privés. À ce stade, il est donc nécessaire pour le moins de comprendre le point de vue, les valeurs et les stratégies des acteurs de chaque niveau de gouvernement.

L'utilité de monographies locales

Dans certains des projets qui ont été soumis au CNE, le problème d'une discordance entre les normes émises par les administrations centrales et leur application sur le terrain a été soulevé. Il en est ainsi du projet sur le logement social dans les DOM et du projet à venir sur l'évaluation de la mise en œuvre des objectifs de sécurité routière au niveau local.

En ce cas, il est nécessaire de comprendre pourquoi certains acteurs locaux ne mettent pas en œuvre une politique, s'il y a des divergences de cultures et d'objectifs entre eux, et si ces divergences sont des effets de statut, de l'histoire, de la non applicabilité pratique des textes, etc. Seules des monographies approfondies permettent de répondre à ces questions et de la sorte aider à ce que les objectifs nationaux soient suivis d'une application effective sur le terrain.

La connaissance du point de vue des usagers

De même il est nécessaire de connaître le point de vue des usagers (par sondages ou enquêtes) qui sont les bénéficiaires finaux des politiques. Sans cette démarche, l'évaluation ne pourra juger de l'impact des politiques publiques, dont seuls les bénéficiaires peuvent témoigner.

Il existe plusieurs méthodes pour connaître le point de vue des usagers et des parties prenantes (celles-ci ont été abondamment analysées par la Brunel University à Londres et le CREDOC en France) :

- sondages de satisfaction ;
- enquêtes de besoin ;
- groupes miroirs participatifs ;
- analyse éthno-méthodologique.

Dans la pratique ces différentes méthodes paraissent s'épauler l'une l'autre.

Les sondages de satisfaction sont les moins satisfaisants car il apparaît dans la plupart des pays que les citoyens sont bienveillants à l'égard de leurs services publics et n'osent pas faire part de certains mécontentements : les sondages de satisfaction traduisent donc rarement les besoins des usagers.

Les enquêtes de besoin sont plus sophistiquées car elles permettent de croiser différents échantillons interrogés avec des variables externes (âge, niveau d'éducation, culture s'il s'agit de différents pays, etc.) et combinent donc l'écoute (entretiens) et des méthodes statistiques.

Les groupes miroirs, c'est-à-dire représentatifs de la diversité des usagers, ont l'avantage de la continuité : un groupe d'usagers apprend à se familiariser avec les contraintes de l'administration ; il y a un apprentissage réciproque, les usagers comprenant mieux l'administration et l'administration comprenant mieux les usagers.

L'éthno-méthodologie (méthode d'observation des rapports agents/usagers sur le lieu de travail) apporte la plus-value de l'observation c'est-à-dire de ce qui se passe réellement dans les relations entre les agents et les usagers.

De plus ces formes de participation sont dynamiques : elles peuvent permettre aux acteurs locaux de tisser des accords sur d'autres manières de faire (il en a été ainsi de l'évaluation de l'habitat social).

3. Les pistes à développer

Le CNE, ne disposant que d'un rapporteur général, a mis l'accent pour sa première année de fonctionnement sur l'analyse des projets venant des administrations et l'élaboration d'un programme à proposer au Premier ministre.

3.1. La participation des élus locaux

Il est apparu difficile à une partie des élus locaux membres du CNE d'en suivre les séances eu égard à leur charge de travail. Ceci ne doit pas apparaître comme un argument opposé à la présence des élus locaux mais au contraire plaider pour que soient choisis des élus moins accaparés et qui ont un investissement personnel dans les pratiques d'évaluation. Le CNE s'est également demandé s'il ne faudrait pas instituer une pratique de suppléance pour les élus locaux désignés membres du CNE.

Les demandes faites au CNE par les collectivités territoriales en matière d'évaluation ont porté plus sur des conseils méthodologiques ou sur des formations plutôt que sur des avis liés à un projet d'évaluation donné. Le sentiment se dégage que les collectivités territoriales reconnaissent la légitimité du CNE pour leur apporter un soutien méthodologique mais moins pour analyser et évaluer « leurs » projets.

Le sentiment du CNE est que les collectivités territoriales ne reconnaissent pas encore le CNE comme un organisme national autonome mais comme un organisme d'État.

Un service de l'évaluation, s'il était créé, pourrait à l'avenir entreprendre une tournée de France de l'évaluation tendant à mobiliser les élus sur ce sujet.

Une base de données Internet sur l'évaluation développant des groupes d'échange et de dialogue, élaborée en conjonction avec la Société française d'évaluation, pourrait contribuer à améliorer la situation.

3.2. La culture d'évaluation dans les administrations publiques

Le CNE a constaté qu'elle avait beaucoup évolué depuis dix ans, certains ministères s'étant dotés de services ou de personnels connaissant bien l'évaluation. Un Club de l'évaluation à la Direction générale de l'administration et de la fonction publique les réunit d'ailleurs régulièrement sur des thèmes d'intérêt commun. Il n'en demeure pas moins qu'il reste un long chemin à faire pour expliquer ce qu'est l'évaluation, quelles en sont les méthodologies et comment les conduire. L'évaluation demeure encore trop liée à l'idée de contrôle ou d'inspection. De plus, nombre d'agents ont fait part de leur souci que l'évaluation des politiques publiques ne mène à l'évaluation des hommes et des femmes qui composent l'administration.

Cette situation supposerait que le CNE en liaison avec le Commissariat général du Plan élabore une sorte de cahier des charges minimum des formations à l'évaluation mais surtout que les cadres, nationaux ou déconcentrés, en perçoivent l'utilité.

La formation à l'évaluation devrait faire partie de toute formation à la modernisation du service public.

S'il y avait un service étoffé de l'évaluation au Commissariat général du Plan, ou au CNE, il serait aussi concevable qu'une base de données des évaluations faites, ou en cours, soit établie et accessible à tous *via* Internet.

Néanmoins des progrès se précisent. Le CNE a été sollicité pour donner des cadrages en matière de formation à l'évaluation par l'ENA (formation initiale et permanente), l'École des Ponts et Chaussées, l'université de Paris-Dauphine, l'Institut international de l'administration publique, certaines IRA. Ces efforts toutefois sont individuels et ne sont pas appuyés par des outils standardisés ni par un site permanent sur Internet.

Le CNE maintiendra pour sa part des liens périodiques avec la Société française d'évaluation et la Société européenne d'évaluation.

3.3 Une nécessaire liaison avec les travaux de la Mission d'évaluation et de contrôle du Parlement

Aujourd'hui, l'évaluation existe à travers de multiples pôles dont la procédure CNE n'est qu'un élément. Il y a les évaluations locales, les évaluations dirigées au sein des ministères, les sujets d'évaluation ad hoc commandités par le

Premier ministre et l'évaluation qui se développe au Parlement en liaison avec la Cour des comptes. Ce foisonnement peut être fécond si les différents participants échangent et évitent de dupliquer leurs évaluations. La montée en force de l'évaluation parlementaire commande au CNE d'établir des liens constructifs avec la Mission d'évaluation et de contrôle du Parlement, d'autant plus que celle-ci a décidé de se concentrer sur le rapport coût/efficacité des administrations, qui pourra être un élément très utile pour les évaluations à venir du CNE.

3.4. Une approche plus internationale de l'évaluation

Les différentes méthodes d'approche de l'évaluation au niveau international sont encore trop peu connues en France et il est important que le débat s'amorce notamment au sein de la CEE.

En effet la comparaison internationale nous permettrait de soulever des questions décisives. Dans certains pays tels que les Pays-Bas et la Suède l'évaluation est fortement poussée et réalisée par les Cours des comptes : quels sont les avantages et les inconvénients de cette formule ? Dans d'autres pays les programmes sont soumis à une obligation d'évaluation à intervalles précis (3 ou 5 ans) ce qui produit de nombreuses évaluations dont certaines ne correspondent qu'à une obligation formelle : que penser de ce système ?

Enfin, dans certains pays, l'évaluation est directement liée au processus budgétaire, non dans une relation mécanique (une bonne évaluation ne donne pas nécessairement plus d'argent et vice versa) mais dans la culture des fonctionnaires de la Direction du budget qui se concentrent non seulement sur l'efficience mais aussi sur l'efficacité : une telle perspective est-elle envisageable en France ?

Le CNE se penchera sur ces questions.

3.5. Une utilisation de la technologie informatique au service de l'évaluation

Un des principaux problèmes rencontrés par les évaluations est celui de l'histoire, tant celle des objectifs des politiques publiques qui ont pu se stratifier dans le temps que celle des recherches et travaux déjà effectués sur ce thème. Aujourd'hui ces reconstructions historiques sont empiriques, en faisant appel à

la mémoire des acteurs, en consultant les services de documentation des administrations ou le service national des thèses, mais il n'existe pas de formule ordonnée permettant de réaliser une bibliographie pertinente.

Certains pays anglo-saxons développent dans les administrations ce qui existe déjà par exemple dans les cabinets d'avocats c'est-à-dire des bases de données informatiques qui par un jeu de mots clés permettent de retrouver tout ce qui s'est écrit sur le sujet en cause, soit directement en fournissant le document soit en donnant sa référence. Cette démarche, appelée *knowledge management* (management du savoir), suppose des systèmes informatiques adéquats mais aussi une collaboration des documentalistes et des gestionnaires pour sélectionner les documents pertinents. Néanmoins un tel système apporte un gain de temps et de mémoire très appréciable.

3.6. Recommandations

En résumé les principales conclusions tirées par le CNE sont les suivantes :

Instances d'évaluation

- Veiller à ce que les instances d'évaluation soient réduites en nombre et non dominées par l'administration ou les groupes ayant le plus intérêt au maintien de la politique évaluée.
- Former les membres des instances aux concepts et méthodes de base de l'évaluation et leur permettre d'établir un dialogue constructif avec le CNE.

L'évaluation

- Favoriser un programme d'évaluation associant des évaluations complexes à des évaluations plus légères, donc plus brèves.
- N'accepter que des projets d'évaluation partant d'hypothèses ciblées sur les politiques à évaluer.
- Demander aux administrations porteuses de projets de reconstituer l'histoire de la stratification des objectifs afin de savoir ce que l'on évalue et de pouvoir se référer à des objectifs.

- Obliger les administrations ou l'instance d'évaluation dès le début d'une évaluation à élaborer des indicateurs de succès (ou d'échec) qui serviront de critères acceptés par les équipes de travail pour mesurer les écarts par rapport aux objectifs.
- Etablir des relations de causalité aussi claires que possibles entre objectifs et résultats.
- Gérer les inquiétudes des parties prenantes de l'évaluation s'il y a des inquiétudes sur la remise en cause de leurs attributions.
- Attendre des instances qu'elles demandent aux équipes d'évaluation d'explorer le rapport coût/efficacité de la politique évaluée.
- Inclure dans les cahiers des charges la nécessité d'une approche internationale.
- Développer partout où cela est utile et possible des monographies locales et des formes d'associations des usagers qui se basent sur l'écoute et le dialogue.
- Mettre en place les moyens de diffuser et développer une culture de l'évaluation dans les administrations.
- Utiliser la technologie informatique au service de l'évaluation.

Concluons par une note positive qui nous est donnée par J. Palumbo ¹ :

«Un processus d'évaluation peut être utile même si l'on ne trouve rien de nouveau. Il peut aider les agents à clarifier ce qu'ils font, à établir des priorités, à mieux allouer leurs ressources et à connaître les domaines qu'ils doivent améliorer».

Plus fondamentalement :

«Le pari de l'évaluation dans un monde imparfait est que les chiffres et les faits peuvent faire une différence, qu'un peu de logique scientifique et d'évidence

(1) « *The Politics of Program Evaluation* », introduction, Sage.

empirique est mieux que leur absence, qu'un accord entre acteurs est mieux qu'une décision arbitraire».

Annexe I

Discours du ministre de la Fonction publique, de la Réforme de l'État et de la Décentralisation

INSTALLATION DU CONSEIL NATIONAL DE L'ÉVALUATION

Commissariat général du Plan
Salle n° 1 - 18 rue de Martignac
Le mardi 16 février 1999 à 17 H 00

ELEMENTS DE DISCOURS

DE M. EMILE ZUCCARELLI,

MINISTRE DE LA FONCTION PUBLIQUE, DE LA REFORME DE L'ÉTAT,
ET DE LA DECENTRALISATION

Seul le discours prononcé fait foi.

Monsieur le Commissaire, Mesdames et Messieurs les membres du Conseil national de l'évaluation, Mesdames et Messieurs,

Par décret du 18 novembre 1998, le gouvernement a créé le Conseil national de l'évaluation pour animer une procédure interministérielle d'évaluation des politiques publiques, plus efficace et plus rapide que le dispositif précédent.

De nombreux commentaires ont salué cet événement. Dans la *Lettre du cadre territorial* de janvier, qui y a consacré un éditorial et une série d'articles fort intéressants, on pouvait lire, entre autres : « **Dans le bowling réglementaire, le pouvoir exécutif vient d'effacer, par une boule précise, le décret du 22 janvier 1990 et de le remplacer par une nouvelle quille, en date du 18 novembre, portant décret relatif à l'évaluation des politiques publiques** ».

Sur un sujet aussi sensible et discuté que l'évaluation, il ne pouvait en aller autrement. Alors que les gouvernements immédiatement précédents l'avaient laissé dépérir, **il y avait urgence à rendre plus opérationnel le dispositif d'évaluation** voulu et mis en place par Michel Rocard pour favoriser la modernisation de l'État **sans rien perdre de son ambition démocratique**.

Votre nomination comme membre du Conseil national de l'évaluation par le Premier ministre, en date du 26 janvier 1999, s'inscrit dans cette perspective. Vous composez

une instance pluraliste, qui allie à un haut niveau de compétence scientifique et juridique, la connaissance et la pratique approfondie des administrations et des institutions et la légitimité démocratique incontestable que lui confère la présence d'élus des collectivités territoriales et de membres du Comité économique et social en son sein.

Au nom du gouvernement, il me revient de procéder à l'installation du Conseil national de l'évaluation. Je rappellerai les grandes lignes de sa mission et les enjeux qui y sont attachés. **Le Conseil procédera ensuite à la désignation de son président, comme il lui appartient de le faire pour achever de se constituer en une instance autonome, entièrement responsable de l'accomplissement de sa mission.**

Les enjeux de l'évaluation des politiques publiques

Le 5 novembre 1997 j'ai fait adopter par le gouvernement les orientations pour la réforme de l'État. L'évaluation des politiques publiques y figure en bonne place, comme instrument de modernisation de la gestion publique et comme modalité de réponse à l'exigence démocratique de transparence et d'efficacité des politiques publiques.

En effet, en dépit des progrès méthodologiques accomplis ces dernières années grâce, notamment, au Conseil scientifique de l'évaluation et du succès d'initiatives prises ici ou là par des services de l'État ou des collectivités locales, la demande d'évaluation reste faible et doit être stimulée. Les pouvoirs publics peinent toujours à rendre compte de leur action, à admettre que celle-ci peut et doit être jugée pour mieux servir l'intérêt général.

L'État est au service des citoyens, sa réforme doit se faire pour eux et avec eux, dans la transparence, le débat et la concertation la plus large. Le temps où la décision publique pouvait se fonder sur le simple exercice de l'autorité est révolu.

En 1998, l'OCDE a étudié les mesures prises par les pays membres pour améliorer l'utilisation de l'évaluation. Ses conclusions mettent l'accent sur la nécessité de gérer les démarches d'évaluation de l'administration publique dans une perspective stratégique de capitalisation des savoirs et de réduction des incertitudes.

Une gestion publique responsable ne peut donc se passer de l'évaluation et il n'est guère de démocratie développée où sa pratique n'ait sensiblement progressé dans les dix dernières années.

Nombreux sont les élus qui se sont inquiétés du devenir du Conseil scientifique de l'évaluation dès l'annonce de mon intention de relancer et d'améliorer le dispositif du décret de 1990. Les régions de Bretagne, du Centre, du Limousin, du Nord-Pas-de-Calais, de Rhône-Alpes ont manifesté leur attachement à une instance nationale à laquelle référer leurs pratiques et qui fasse autorité à l'étranger.

Je me suis engagé à ne pas décevoir cette attente, d'autant que la réussite et le développement de la coopération administrative internationale l'exigent. Il y va en effet de la crédibilité de nos méthodes et de notre offre de formations. **Le maintien de**

l'influence française exige que nous soyons présents de manière générale sur le terrain de l'évaluation ce qui, compte tenu de la faiblesse du milieu spécialisé, passe par l'existence d'un pôle fort et reconnu.

Le Conseil scientifique de l'évaluation lui-même n'avait pas manqué de s'interroger sur le bien-fondé d'une relance de l'évaluation au niveau gouvernemental. Dans son *Rapport d'activité pour l'année 1997*, il nous a prévenu de deux écueils :

Le premier serait de « magnifier » l'évaluation au point qu'elle paraisse réservée à l'analyse des « grandes politiques » et de décourager ipso facto toute velléité d'appropriation de l'évaluation par les services dans leur activité quotidienne.

Le second écueil serait d'imposer à nouveau l'évaluation plutôt que d'en faire naître le besoin. Dans ce cas, il y a de forte chance pour que l'intérêt porté à l'évaluation ne dure guère plus que l'impulsion donnée. C'est ainsi que doit s'analyser l'échec relatif du dispositif de 1990. D'autres motifs y ont contribué. On peut citer :

- la complexité du processus de sélection des sujets qui comportait 7 étapes successives ;
- l'existence de conflits entre administrations sur le choix des politiques à évaluer, ce qui a favorisé l'émergence de sujets périphériques ;
- les difficultés de pilotage et de réalisation des évaluations résultant de la sous estimation du nombre des experts à réunir sur un projet déterminé, de la complexité des méthodes choisies et de la longueur des délais de publication ;
- l'absence d'un instrument interministériel actif et puissant d'insertion des conclusions d'évaluation dans les processus de décision relatifs aux politiques publiques évaluées.

De sorte, qu'hormis le travail méthodologique du Conseil scientifique de l'évaluation, le bilan global de ce dispositif s'est révélé décevant. Quatorze évaluations, seulement, auront été réalisées en sept ans. Les sujets traités sont d'intérêt inégal en termes de décision publique.

Il faut reconnaître que certains rapports manquent de concret, que leur qualité méthodologique laisse souvent à désirer et qu'ils ont été suivis de peu d'effets. **Toutes choses que le Conseil scientifique de l'évaluation a lui-même soulignées.**

Il nous fallait donc revenir à l'ambition affichée par le rapport Viveret au Premier ministre en 1989 : « **Faire de l'évaluation un levier essentiel de la transformation et de la modernisation de notre système public** », dans le respect des principes de transparence, de pluralisme, de rigueur et d'indépendance affirmés entre temps par le Conseil scientifique de l'évaluation.

Que trouvons-nous par ailleurs dans le champ de l'évaluation ? Des structures qui ont des champs de compétence ou des moyens plus réduits ainsi que des initiatives ministérielles récentes.

La Cour des comptes a développé ses capacités d'évaluation en privilégiant les questions touchant à la gestion et aux finances publiques. Elle y consacre des moyens peu importants et associe peu de partenaires extérieurs à la réalisation de ses travaux, comme des universitaires, des économistes, des sociologues ou des statisticiens.

Le Comité central d'enquête sur le coût et le rendement des services publics est un organisme tripartite - administration, élus, syndicats - placé sous l'autorité du Premier ministre et présidé par le Premier président de la Cour des comptes. Celui-ci m'a récemment indiqué vouloir accroître le rythme de ses travaux. Mais là aussi, le fonctionnement des services publics ne couvre qu'une partie du champ de l'évaluation des politiques publiques.

S'il est encore trop tôt pour juger du fonctionnement des offices parlementaires d'évaluation des politiques publiques et de la législation, on observera que ceux-ci ont une vocation uniquement parlementaire.

Les autres structures, telles que la Direction en charge de l'évaluation au ministère de l'Éducation nationale, le Conseil national d'évaluation de la recherche, le Conseil national d'évaluation de l'enseignement supérieur et de la recherche, le Comité national d'évaluation des établissements publics à caractère scientifique, culturel et professionnel ou l'Agence nationale d'accréditation et d'évaluation en santé, se cantonnent le plus souvent à un secteur technique particulier.

Enfin, par circulaire du 3 juin 1998 relative à l'élaboration des programmes pluriannuels de modernisation des administrations, le Premier ministre a demandé aux ministres de renforcer leur capacité d'évaluer les politiques dont ils ont la charge. La Culture, l'Équipement, l'Agriculture, l'Emploi et la Solidarité et la Défense se sont ainsi dotés de services ou de comités ministériels ad hoc.

L'étape suivante devrait concerner les échelons déconcentrés dont le rôle dans la mise en œuvre et l'articulation des politiques publiques va croissant. C'est en effet sur le terrain que se développent conjointement et s'organisent les politiques de l'État, des collectivités locales et, de plus en plus, de l'Union européenne. L'évaluation des effets économiques et sociaux qui en résultent présente un degré de complexité nouveau auquel les services ne sont pas préparés. Des initiatives doivent être prises rapidement pour surmonter cette difficulté.

Le champ de l'évaluation prévu par le décret du 18 novembre 1998 a été élargi en conséquence. L'objectif est désormais d'apprécier l'efficacité des politiques publiques conduites par l'État, les collectivités locales et leurs établissements publics.

le mandat du Conseil

Je ne reviendrai pas sur les modalités de l'évaluation. Elles sont nombreuses et il vous appartiendra d'en décider. Je m'en tiendrai à l'approche consistant à définir l'évaluation

comme une démarche, un mode de raisonnement systématique permettant d'articuler rigoureusement les jugements portés aux faits analysés.

Le Conseil national de l'Évaluation a pour mandat de proposer au Premier ministre un programme annuel d'évaluation et d'en suivre la mise en œuvre.

À cet effet, il expertise et sélectionne les projets qui lui sont transmis par les ministres, le Conseil d'État, la Cour des Comptes, le Médiateur de la République, le Conseil économique et social, les collectivités territoriales et les associations d'élus : Association des maires de France, Association des présidents de conseils généraux et Association des régions de France. **Si nécessaire, il y ajoute ses propres propositions.**

Le Conseil national de l'Évaluation formule aussi des recommandations à l'attention du gouvernement pour orienter le développement des connaissances sur l'efficacité des politiques publiques et à l'attention de la communauté scientifique pour faire progresser les normes professionnelles et méthodologiques de l'évaluation.

Il effectue sa mission dans une transparence totale. Le programme arrêté par le Premier ministre est publié au *Journal Officiel de la République française*. Il expose le contenu des projets retenus et les raisons justifiant ce choix. Le rapport annuel du Conseil et l'ensemble de ses avis sur les travaux effectués sont également rendus publics, de même que les rapports d'évaluation et l'avis des autorités concernées sur les suites à leur donner.

Le Commissariat général du Plan se voit parallèlement confier le développement de l'évaluation dans les administrations et la mise au point des suites à donner aux évaluations qui concernent l'État.

Assisté dans son fonctionnement par un rapporteur général et par le Commissariat général du Plan, qui en assure le secrétariat, **le Conseil national de l'Évaluation est le moteur de la nouvelle procédure interministérielle d'évaluation des politiques publiques. Il en fixe les échéances et les modalités dans le cadre des dispositions fixées par le décret du 18 novembre 1998 et la circulaire du Premier ministre du 28 décembre relative à l'évaluation des politiques publiques dont j'ai obtenu la publication au *Journal Officiel* du 12 février.**

Sachant que vous serez appelés à en examiner le détail dans quelques minutes, je vous formulerai maintenant les principaux objectifs qu'il vous faudra tenir.

En règle générale, les évaluations que vous sélectionnerez devront permettre d'aboutir dans des délais brefs à des conclusions précises et concrètes, rapidement utilisables dans le cadre des processus de décision qui conduisent à lancer, compléter, infléchir, suspendre ou abandonner une politique publique.

Vous veillerez plus particulièrement à la qualité de l'élaboration des projets pour encourager les autorités et les services publics à s'engager dans des processus d'apprentissage efficaces à tous les niveaux d'administration.

À terme, le meilleur indicateur de votre réussite sera le niveau de déconcentration duquel les évaluations seront couramment effectuées.

La création et l'installation du Conseil national de l'évaluation répondent à l'ambition initiale de disposer d'une instance à même de faire progresser et de garantir les règles déontologiques et leurs méthodes de l'évaluation.

Cette fonction, qui doit être assurée en toute impartialité est essentielle pour assurer la crédibilité des évaluations aux yeux des médias et de l'opinion ainsi que des agents et des décideurs des services publics.

Il vous appartiendra donc de consacrer une partie de vos moyens à communiquer pour faire connaître vos travaux.

De même, je vous demande d'établir toutes relations utiles avec les instances d'évaluation qui travaillent dans les autres pays pour permettre à la France de jouer un rôle dans ce domaine fortement marqué par les échanges entre experts internationaux et jusqu'à présent largement dominé par les experts anglo-saxons.

La publication de vos avis et rapports sur Internet et en langue anglaise serait la bienvenue à cet effet.

Enfin, tout en marquant davantage les préoccupations légitimes de la collectivité nationale en matière de productivité, de qualité et de coût des politiques publiques, vous veillerez à ce que les préoccupations d'ordre économique soient prises en compte sans perdre de vue l'intérêt et les nécessités d'une approche qualitative et sociologique des effets des politiques évaluées.

En tant que ministre chargé de la Réforme de l'État, j'attends du Conseil national de l'évaluation qu'il fasse autorité dans son domaine et que, guidé par une haute conception de l'action publique, il contribue à structurer efficacement la démarche d'évaluation des pouvoirs publics.

L'évaluation est un instrument fondamental de la Réforme de l'Etat. Je sais pouvoir compter sur vous pour en faire ressortir l'intérêt et les avantages et la promouvoir. Je vous en remercie.

Monsieur le Commissaire, Mesdames et Messieurs les membres du Conseil national de l'évaluation, Mesdames et Messieurs, le Conseil national de l'évaluation est installé.

Je vous remercie de votre attention. M. Edouard Salustro, je vous cède la présidence de séance avant de vous laisser procéder à la désignation du président du Conseil national de l'évaluation.

Annexe II

Lettre du CNE au Premier ministre

Conseil national de l'évaluation

PARIS, le 29 juin 1999

Réf. : PROJET DE PROGRAMME 1999

Affaire suivie par : YC/ST

Monsieur le Premier ministre,

Conformément à l'article 4 du décret du 18 novembre 1998 relatif à l'évaluation des politiques publiques, le Conseil national de l'évaluation, installé le 16 février 1999, vous propose son premier programme d'évaluation établi en étroite collaboration avec le Commissariat général du Plan.

Compte tenu des délais de formulation des projets par les ministères puis de leur instruction, ce programme comprend deux groupes :

- Six projets sont prêts à être lancés. Les évaluations pourront commencer aussitôt que vous les aurez approuvées et que les instances auront été mises en place. Il s'agit de :
 - les effets de la politique agricole commune (PAC) sur l'environnement ;

- la politique de préservation des ressources destinées à la production d'eau potable ;
 - les aides à l'emploi dans le secteur non marchand ;
 - le programme « nouveaux services-emplois jeunes », dans le champ d'intervention du ministère de la Jeunesse et des Sports ;
 - la politique de lutte contre le SIDA ;
 - la politique en faveur du logement social dans les départements d'Outre-mer.
- Nous avons déjà été saisis d'autres propositions, à des degrés divers d'avancement, qui devront être plus précisément examinées, de concert avec les administrations concernées, pour une sélection d'ici l'automne. À ce stade, leur liste est seulement indicative et ne comporte aucun ordre de priorité. Il s'agit de :
 - la fonction juridique dans les services déconcentrés ;
 - les aides à la presse ;
 - la formation des agents de l'État ;
 - la concession de service public ;
 - la coopération administrative internationale ;
 - la sécurité routière ;
 - le logement des agents de la Défense ;
 - les Fonds structurels européens.

L'analyse détaillée des six projets sélectionnés a été conduite selon les critères suivants : intérêt général et opportunité du thème ; faisabilité de l'évaluation (coûts et délais) ; méthodes et problèmes à prendre en compte ; pertinence du profil proposé de l'instance.

Conformément aux orientations que vous avez arrêtées, ce programme privilégie une approche interministérielle, qui devra souvent être élargie aux collectivités locales par les instances d'évaluation à désigner. Il concerne des domaines importants comme l'emploi, l'environnement, l'agriculture, la santé, le logement, et s'attache aux priorités fixées en matière d'emploi, d'environnement, de compétitivité pour la France et de qualité et sécurité des services rendus aux usagers.

Les projets choisis marquent la volonté des administrations de s'inscrire dans une hiérarchie des priorités d'évaluation pleinement cohérente avec les thèmes majeurs de l'action du gouvernement et de sa politique de modernisation du service public. Mais ces projets ne sauraient représenter ou résumer le programme d'évaluation dont vous avez fixé l'ambition en établissant le Conseil.

C'est ainsi que si la place faite aux préoccupations d'environnement est centrale – puisque parmi les six premiers figurent deux projets liés à l'environnement – il ne s'agit toutefois que d'un premier pas sur le chemin d'une appréhension évaluative de maintes politiques publiques dans leurs effets sur l'environnement.

C'est ainsi que, pour l'heure, le Conseil déplore l'absence de propositions portant sur des politiques économiques ou des enjeux budgétaires plus notables, et regrette entre autres que les politiques d'achats publics et de commande n'aient pas été évoquées. Le social, hormis les politiques publiques d'éducation nationale, semble plus évalué que les mécanismes de régulation économiques et juridiques qui régissent notre quotidien et appellent des comparaisons avec nos partenaires européens.

Le Conseil, en liaison étroite avec le Commissariat général du Plan, s'emploiera donc à rééquilibrer dans ce sens les projets d'évaluation à venir, en les inscrivant dans une perspective à moyen terme, en veillant à ce que l'approche interministérielle soit concertée entre les ministères concernés comme avec les autres acteurs publics impliqués, et que chacun des projets traduise des préoccupations essentielles au gouvernement.

Le Conseil s'est soucié de réconcilier l'expertise et les impératifs de l'action administrative en vous proposant à la fois des évaluations aux enjeux complexes et des évaluations qui pourront être menées de façon plus rapide en un an. Afin que ces évaluations soient constructives et tirent les leçons du passé pour une meilleure gestion de l'avenir, le Conseil a privilégié la mise au point d'outils et de techniques qui faciliteront l'action administrative.

Certains aspects demanderont à être précisés ou amendés lors du lancement des évaluations :

- Les instances d'évaluation proposées tendent toujours à être trop larges par souci d'inclure tous les acteurs de la politique publique mais avec le danger de paralyser le pilotage effectif des travaux. Le Conseil propose des instances plus réduites qui s'entourent d'auditions ou de groupes de travail. Le recrutement de leurs membres devrait s'opérer, à l'initiative du Commissariat général du Plan, secrétaire du Conseil, sur des critères tels que : la compétence, l'absence de conflit d'intérêt, la disponibilité et la présence *intuitu personae*. La personnalité du président est décisive : son objectivité et son indépendance seront tout aussi déterminantes que ses capacités d'écoute et de synthèse ou de maîtrise du calendrier tendu des évaluations.

- Compte tenu de la charge de travail qui incombera aux instances et de la nécessité d'en stabiliser les membres mieux que par le passé, la question de la rémunération de leurs membres se posera. De telles indemnités existent dans d'autres pays de l'Union européenne et sont partie intégrante du financement des travaux d'évaluation.
- Le questionnement proposé par les administrations est souvent trop riche pour permettre leur évaluation dans les délais impartis. Dans la plupart des projets, il serait bon que les questions soient groupées en quatre ou cinq hypothèses fondamentales qu'il s'agirait de vérifier. Tant le Commissariat général du Plan que le CNE se sont employés à procéder à ce regroupement qui figure dans les fiches qui vous sont jointes sur chaque projet d'évaluation.
- Les administrations gestionnaires ont relativement peu d'expérience des méthodes de l'évaluation. L'aide des rapporteurs du CNE et des experts du Commissariat général du Plan devra être mobilisée, mais peut-être aussi celle de consultants. En outre, les instances seront incitées à recenser et tirer le meilleur parti des études et travaux existants, français et internationaux, avant d'en engager de nouveaux.

Enfin le Conseil tient à saluer l'intérêt actif des administrations et les efforts importants que celles-ci ont déployés dans l'élaboration des projets construits dans des délais très rapides.

Cette montée en régime de l'évaluation se poursuivra, dès votre décision sur ce programme, d'abord par le choix et la mise en œuvre des instances, projet par projet, au cours de l'été. C'est avec les présidents qui seront choisis pour animer les instances que le Conseil et le Commissariat s'emploieront à préciser leur mode de composition, les critères de choix des opérateurs publics ou privés chargés de l'évaluation, et le calendrier des travaux envisagés. Ensuite, à l'automne, le Conseil sélectionnera un deuxième groupe de projets pour le soumettre à votre décision avant la fin de l'année.

Yves Cousquer

Annexes à la lettre
Projet de lutte contre le SIDA
Intérêt général et opportunité

L'intérêt politique de ce projet est majeur car il aborde les questions de fond relatives à une maladie très sensible en matière d'opinion publique et cherche à trouver des solutions constructives pour la gérer socialement, en termes de solidarité et de bonne utilisation des deniers publics.

Il prend en compte la montée en charge des associations dans leur action bénévole travaillant avec les services de l'État et les collectivités territoriales.

Il promeut un modèle de l'action publique déconcentrée où les synergies s'établissent entre les acteurs locaux qu'ils soient d'État, des collectivités territoriales ou du monde associatif.

Le point le plus important de ce projet est de souligner l'évolution épidémiologique et thérapeutique de la maladie qui a complètement changé, et peut maintenant faire appel à des dispositifs suivant des malades sur des dizaines d'années.

Faisabilité

La principale recommandation à ce titre est de déterminer ce qui est développé dans le cadre spécifique de la prévention et ce qui participe au traitement du SIDA aujourd'hui face aux évolutions en cours de la maladie dans les dispositifs publics.

Méthodes et suggestions

Certaines grandes questions pourraient aider à structurer le projet :

L'évaluation du rôle de la solidarité

- Quel est le rôle de la solidarité en matière de prévention et de traitement ? Cette solidarité qui est déterminante est-elle correctement mise en œuvre et comment peut-elle être améliorée ?
- Dans la mesure où cette action se situe surtout au niveau local, il serait bon d'établir des monographies qui éclairent les logiques des acteurs publics et

associatifs et en tire des leçons pour améliorer les actions, les collaborations et le travail en commun.

La nécessité d'un système de mesure

- Une analyse et une mesure des ressources allouées et de leur impact afin de tester si le rapport coût/efficacité est bon.
- Une évaluation des moyens nécessaires pour mettre en place des bases de données et une mesure utilisable en temps réel de statistiques sur le SIDA ou la séropositivité qui soient utiles aux décideurs.

Une évaluation de l'effectivité de la prévention

- Cette analyse sera replacée dans un contexte international où les choix en matière de prévention ont été fort différents.

L'instance d'évaluation

- Elle devra être plus ramassée, être dotée d'un président reconnu pour son impartialité en la matière et d'un rapporteur spécialiste des méthodes d'évaluation. Bien entendu, cette instance pourra s'entourer de toute l'expertise nécessaire soit à travers des auditions ou des groupes de travail ponctuels.

L'évaluation des "aides à l'emploi dans le secteur non marchand" et le projet d'évaluation "nouveaux services-emplois jeunes"

Il s'agit de deux projets différents, présentés respectivement par le ministère de l'Emploi et de la Solidarité et par le ministère de la Jeunesse et des Sports. Compte tenu du fait qu'ils recouvrent pour partie le même champ, il semble opportun d'en faire une présentation conjointe et de souligner la nécessité de veiller à leur complémentarité.

Les emplois aidés dans le secteur non marchand

Intérêt général et opportunité

Le premier projet concerne l'ensemble des dispositifs d'aide à l'emploi dans le secteur non marchand mis en œuvre depuis le milieu des années 1980 : travaux d'utilité collective (TUC), programmes d'initiative locale (PIL), remplacés en 1990 par les contrats emploi –solidarité, les emplois ville et les contrats emplois consolidés (CEC). Le programme « Nouveaux services– nouveaux

emplois » (les emplois-jeunes), en dépit de son caractère innovant, s'inscrit dans la continuité de cette politique dont il constitue le plus récent développement. L'intérêt d'évaluer cette politique résulte d'abord son importance quantitative (444 000 personnes étaient titulaires d'un contrat de travail aidé dans le secteur non marchand à la fin de l'année 1998). L'intérêt du projet résulte également du fait que les mesures à évaluer visent pour l'essentiel les publics dont l'accès à l'emploi est le plus difficile, et qui bénéficient le moins de la reprise actuelle de l'emploi. Parallèlement à cet objectif d'insertion, les dispositifs d'aide à la création d'emploi dans le secteur non marchand ont également pour finalité, et cela depuis la création des TUC, de favoriser le développement d'activités nouvelles répondant à des besoins insuffisamment pris en compte par les services publics et les entreprises privées.

Sur fond de permanence de ces deux objectifs, l'évolution des dispositifs a été marquée par un mouvement de pérennisation et de consolidation des statuts d'emploi, avec la création des CEC puis des emplois-jeunes, ces derniers se traduisant en outre par une nouvelle segmentation au sein du public cible.

Faisabilité

Les questions posées par le projet d'évaluation élaboré par le ministère de l'Emploi et de la Solidarité concernent classiquement l'effet net de cette politique sur l'emploi et le chômage, l'efficacité des dispositifs en terme de réinsertion professionnelle des bénéficiaires, leur impact sur la gestion de l'emploi et de développement de l'activité dans les organismes employeurs, l'apparition de nouveaux métiers et de nouvelles compétences, la révélation d'une nouvelle demande de service, solvable ou non.

Méthodes et suggestions

Compte tenu du nombre et de la qualité des études déjà réalisées relativement à certaines de ces questions, le Conseil national de l'évaluation recommande d'en effectuer un bilan précis, et de réserver les nouvelles études aux points qui ont été jusqu'ici insuffisamment étudiés :

- Quel est l'impact à long terme du passage en emploi aidé sur les parcours professionnels des bénéficiaires ?
- Est-il possible d'analyser en terme économique la valeur créée par les activités développées grâce aux emplois aidés ?
- Quel est l'impact de ces emplois sur la gestion des ressources humaines dans le secteur public (effets de substitution, amorce d'une nouvelle segmentation

ou mise en place d'une nouvelle modalité de recrutement des agents publics) ?

- La gestion des dispositifs par le service public de l'emploi est-elle favorable à la mobilisation du secteur associatif ?

Ces différentes questions valent pour l'ensemble des dispositifs considérés. Cependant, l'évaluation devra également prendre en compte et analyser les innovations, voire les ruptures, introduites par le programme « Nouveaux services–Nouveaux emplois », du point de vue notamment des préoccupations suivantes :

- La priorité accordée au développement d'activités répondant à une véritable demande sociale potentiellement solvable.
- Le renforcement d'une logique de professionnalisation (l'identification de nouveaux métiers).
- La volonté de passer d'une logique de guichet à une logique de projet, instaurant une nouvelle forme de coopération entre l'administration et le secteur associatif.

Instance d'évaluation

S'agissant des modalités de pilotage de cette évaluation, il semble préférable d'opter pour une instance d'évaluation de taille réduite composée de personnes disponibles et compétentes, disposées à s'investir fortement pendant toute la durée de l'évaluation. Quitte pour cette instance à procéder à des auditions, ou à s'élargir ponctuellement à un groupe plus nombreux comprenant l'ensemble des partenaires administratifs ou non susceptible d'avoir un point de vue pertinent sur les questions étudiées.

Le programme “nouveaux services - nouveaux emplois” dans le champ d'intervention du ministère de la Jeunesse et des Sports

Ce projet, élaboré par le ministère de la Jeunesse et des Sports, concerne le programme « nouveaux services–nouveaux emplois », dans le seul domaine de compétence du ministère. Les principales questions relatives à l'évaluation de l'ensemble des emplois aidés dans le secteur non marchand se posent également dans ce champ plus restreint. Il semble donc justifié de considérer ce second projet comme un approfondissement du précédent sur un domaine et un segment particuliers de la politique. Cela signifie que l'évaluation devra s'attacher plus particulièrement à l'analyse des évolutions induites par le programme dans le champ d'activité de la Jeunesse et des Sports (en termes de

nature des activités développées grâce au programme, d'impact sur les compétences, les formations et les diplômes, d'impact sur le tissu associatif et les relations entre ce dernier et les administrations, ainsi que de partenariat inter-administratif).

Le choix de ce domaine particulier de l'action publique pour évaluer de manière approfondie les effets du programme « nouveaux services–nouveaux emplois » semble pertinent, compte tenu de la dynamique quantitative qui s'y dessine (30 000 emplois créés au 1^{er} janvier 1999), et du rôle joué par les associations (60 % des emplois créés l'ont été par des associations), mais il convient dès lors d'observer que toutes les conclusions de cette évaluation ne seront pas nécessairement transposables à d'autres secteurs utilisateurs des emplois–jeunes. Il importe par ailleurs de tenir compte du fait que certaines questions parmi les plus importantes - celles qui concernent le devenir professionnel des jeunes à l'issue des contrats -, ne pourront être valablement traitées que dans plusieurs années. Il convient donc de considérer le projet d'évaluation figurant dans le présent programme comme la première étape d'un projet de plus longue échéance qui cherchera à mesurer et analyser *ex-post* les effets du programme sur l'emploi et les parcours professionnels des bénéficiaires.

Enfin, dans le but d'assurer une bonne coordination entre les deux évaluations relatives aux emplois aidés dans le secteur non marchand (et notamment une complémentarité des programmes d'étude), il convient qu'au minimum une personne soit membre des deux instances d'évaluations chargées de les piloter.

Projet d'évaluation de la "politique de préservation de la ressource destinée à la production d'eau potable"

Intérêt général et opportunité

Ce projet présente un intérêt certain car il recouvre une évolution de fond qui est le passage de l'idée de l'eau comme ressource illimitée à l'eau comme un bien rare. L'évaluation de « la politique de préservation de la ressource destinée à l'eau potable » répond à des nécessités politiques, environnementales et de santé publique mais aussi à des préoccupations économiques.

- Nécessité politique double du fait de la préoccupation croissante des consommateurs français (marquée par une extension du marché des eaux de source) et à des recours en responsabilité de plus en plus fréquents auprès des instances européennes pour non-conformité des eaux françaises à la réglementation européenne.

- Nécessité de santé publique car la dégradation des eaux de surface s'accélère comme celle des eaux profondes. Les moyens de mesure de ces dégradations semblent aujourd'hui défectueux.

Faisabilité

- L'évaluation proposée est étayée par un dossier déjà très solide reposant sur des études préalables (Commissariat général du Plan, Cour des comptes) de qualité.

Méthodes et suggestions

Le CNE a alors suggéré de structurer le questionnement.

L'analyse de l'effectivité des dispositifs juridiques

Le projet souligne l'abondance des dispositifs juridiques existants et leur non-application sur le terrain, notamment parce que les acteurs locaux d'État ou les collectivités territoriales se sentent tenus de faire prévaloir des logiques économiques à court terme ou ne se préoccupent pas de la mise en œuvre effective d'instruments juridiques.

Cette évaluation prendra en considération le dispositif existant, les difficultés d'application qu'il rencontre et ses contradictions internes.

Les solutions alternatives à la politique actuelle, centrée sur la réparation et l'opportunité d'actions de préservation en amont au caractère préventif affirmé

L'évaluation des évolutions sociologiques et économiques

L'évaluation aidera à prendre la mesure du problème en démontrant précisément s'il s'agit de cas marginaux ou d'un problème national et quelles pourraient en être les conséquences économiques (étude économique).

L'évaluation s'attachera à l'analyse des changements comportementaux des français à l'égard du problème (plaintes déposées, développements consuméristes).

L'évaluation mesurera des ressources utilisées y compris le temps des fonctionnaires (étude comptable et budgétaire).

L'évaluation se demandera quels sont les enjeux économiques internationaux de la politique de l'eau.

Le système d'incitations/sanctions

- La logique du pollueur-payeur semble en faire plus un pollueur sociétaire : un autre mode d'incitation et de sanction pourrait-il être imaginé ?
- Les logiques contractuelles avec les agriculteurs portent-elles leurs fruits ?
- Y a-t-il des blocages culturels et financiers au niveau des corps de l'État ? Ces blocages ou ces ignorances peuvent-ils être levés et comment ?

• Face aux méthodes, le CNE a fait plusieurs suggestions :

1. Etablir un échantillonnage de zones qui feraient l'objet d'une étude monographique auprès de l'ensemble des acteurs pour comprendre le pourquoi de la non-application des textes (analyse sociologique).
2. Ces zones devraient être choisies en fonction des critères déterminants de la politique élaborés par le ministère et mélanger « bons » et « mauvais » exemples ainsi que les zones préservées et les zones à reconquérir.
3. La dimension comparative est essentielle : existe-t-il une comparaison effectuée au niveau de l'Union européenne ou d'autres pays (étude comparative) ?

Instance d'évaluation

Comme dans les autres projets le CNE suggère que l'instance soit plus ramassée et conduite par un président motivé dont l'ouverture internationale est connue et qui ait une faculté d'écoute établie dans la diversité des points de vue et intérêts engagés par cette politique.

L'évaluation des "effets de la PAC sur l'environnement"

Intérêt général et opportunité

Ce projet présente un degré d'opportunité élevé car il touche un problème directement sensible aux citoyens, celui de l'environnement et répond au cinquième programme communautaire d'action sur l'environnement 1993-2000.

L'intérêt politique de cette évaluation est fort dans la mesure où la France est engagée dans des négociations internationales complexes où d'autres pays ont constitué des blocs de position unifiée (groupe de Cairns, Uruguay round). Les prochaines négociations sur les échanges en 2002 réviseront les systèmes de soutien à l'agriculture.

L'évaluation portera sur la politique socio-structurelle (en particulier les mesures agro-environnementales) et la politique des marchés (céréales et oléoprotéagineux, viande bovine, lait). Ce champ n'est pas trop vaste car ces différentes filières sont intriquées, de plus il ne serait pas possible d'évaluer rétrospectivement l'ensemble de la politique agricole française.

Faisabilité

La qualité de ce projet d'évaluation est à souligner :

- La politique, ses objectifs et leur évolution dans le temps sont précisément analysés.
- Les problèmes sont clairement mis en évidence, notamment :
 1. Une vision unitariste qui ne prend pas en compte les différences entre les pays ;
 2. La perte de maîtrise d'un système visant à l'accroissement de la production ;
 3. L'incapacité des volets structurels à éviter la désertification ;
 4. Un constat à travers les études et chiffres existants d'un bilan dommageable à l'environnement ;
 5. La complexité des rapports entre économie, agriculture et environnement.
- Les résultats attendus de l'évaluation sont définis :
 1. Une base d'amélioration des objectifs environnementaux intégrés aux politiques publiques agricoles, notamment la notion de bilan environnemental, la construction d'une position de négociation internationale et une pédagogie des acteurs.

2. La mise en place de garde-fous adaptés via la conditionnalité des aides, la recherche de solutions techniques appropriées, par la contractualisation pluriannuelle de programmes agri-environnementaux, la recherche de moyens de maîtrise de la concurrence pour éviter la disparition des systèmes agricoles favorables à l'environnement.

Méthodes et suggestions

Le projet propose plusieurs approches que le CNE soutient :

- Une mesure de l'efficacité de la PAC dans ses effets recherchés et non recherchés. Cette mesure suppose que la bibliographie indicative du projet soit complétée par des travaux européens et régionaux. La dimension comparative i.e. les choix faits par d'autres pays devrait servir de guide (Pays-Bas, Autriche).
- Une analyse économique de l'intégration des marchés et des marges de manœuvre disponibles.
- Une analyse sociologique du comportement des acteurs. À ce titre le point de vue des consommateurs sera étudié ainsi que celui des organisations syndicales, des coopératives et de la Mutualité sociale agricole.
- Une analyse du lien entre les modes de soutien économiques et la mobilisation des moyens de production.
- Une vision prospective de l'impact de la PAC sur l'environnement par la construction d'indicateurs qui permettront de suivre cette politique pour le futur.

L'instance d'évaluation

Comme pour les autres projets celle-ci devra être plus compacte et éviter des membres ayant des conflits d'intérêt.

La réorientation vers le préventif est potentiellement source d'économies budgétaires sensibles et de meilleure efficacité.

L'évaluation de la politique " interministérielle en faveur du logement social dans les départements d'Outre-mer"

Intérêt général et opportunité

Ce projet a été jugé spécialement pertinent au regard de l'intérêt général dans la mesure où les questions relatives aux DOM sont très sensibles pour l'évolution du territoire français et de la cohésion nationale.

Il semble que le logement social dans les DOM ait fait jusqu'à présent l'impasse sur l'insertion dans une politique urbaine et environnementale, sur les besoins des habitants, les contraintes du foncier et ait plaqué des solutions décalquées de celles de la métropole sur des territoires qui ne l'étaient pas.

Faisabilité

Une inquiétude s'est faite jour car nombre de questions de mesure figurant dans le projet d'évaluation devraient être d'ores et déjà connues et utilisées par les administrations au jour le jour. Une évaluation ne peut se substituer au contrôle de gestion. Il serait bon que l'évaluation soit précédée par la mise à jour des données disponibles et des analyses qualitatives des points de vue des acteurs de terrain, dont les préfets.

Méthodes et suggestions

Les observations du CNE tendent à pousser le projet vers des questions plus fondamentales et structurantes :

La clarification des objectifs de la politique publique et de son impact

- Tels qu'ils sont formulés, les objectifs de la politique sont ceux d'un opérateur urbain, ayant à équilibrer ses comptes. Le CNE propose l'interrogation suivante acceptée par les administrations compétentes : quelle est la politique publique réelle de logement social dans des territoires à fort taux de chômage et à RMI très élevé et quelle devrait être la politique souhaitable ?
- Quels sont les effets pervers de cette politique : profite-t-elle à ceux qui en ont le plus besoin ou à ceux qui ont déjà des ressources ?
- Y a-t-il adéquation avec les besoins de la population et si non quelles pourraient être les solutions alternatives ?
-

La précision de l'approche méthodologique

- Le projet est en cours de développement d'une approche méthodologique. De ce fait le CNE propose une approche économique qui analyse la viabilité de la politique de logements sociaux au regard des données économiques des DOM et de leurs habitants ; le CNE suggère aussi une approche sociologique des besoins par divers moyens d'enquête qui pourrait s'appuyer sur ce qui a été fait lors de l'évaluation du logement social au niveau national. Ces deux propositions s'intègrent dans la démarche d'évaluation telle que les administrations la conçoivent.

L'instance d'évaluation

- Il paraît enfin indispensable que l'instance nationale soit relayée par des instances locales et pas seulement par des experts en visite.

Annexe III

Evaluations en cours

Évaluation de la politique de préservation de la ressource en eau destinée à la consommation humaine

Président : Frank Villey-Desmeseretz
Rapporteur général : Denis Ballay
Rapporteurs adjoints : Dominique Tricard – Caroline de Villeneuve

Évaluation du programme « nouveaux services–emplois jeunes » dans le champ d'intervention du ministère de la Jeunesse et des Sports

Président : Anicet Le Pors
Rapporteur général : Jean-Pierre Chassiné
Rapporteur adjoint : François Boddaert

Évaluation des mesures d'aide aux emplois du secteur non marchand

Président : Yves Robineau
Rapporteur général : Bernard Simonin
Rapporteur adjoint : Guy Truchot

Évaluation du programme 1995-2000 de politique de lutte contre le SIDA

Président : Pierre Ducimetière
Rapporteur général : Michel Duraffourg
Rapporteurs adjoints : Claudine Parayre – Gérard Pelé
Évaluation de la politique du logement social dans les DOM

Président : Jean-Luc Mathieu
Rapporteur général : Jean-Marie Butikofer
Rapporteuse adjointe : Corinne Desforges

Contribution I

Contributions au séminaire du CNE du 25 novembre 1999

1. Quelques concepts clés pour définir un projet d'évaluation par Eric Monnier (C3E¹)

Ce texte reprend de larges extraits du volume n°1 de la Collection MEANS que nous avons préparé pour la Commission européenne : « Évaluer les programmes socio-économiques. Conception et conduite d'une évaluation », Luxembourg, Office des publications officielles des Communautés européennes, 1999.

L'évaluation prend de multiples formes selon le stade où elle intervient dans le cycle de programmation (évaluation ex-ante, mi-parcours, ex-post), selon le niveau de décision concerné (évaluation de politique, de programme, de projet), et selon son caractère plus ou moins étendu (évaluation d'ensemble, évaluation thématique, évaluation plus approfondie de certains aspects d'une politique ou d'un programme). Loin d'être indépendantes les unes des autres, ces différentes formes d'évaluation doivent s'articuler les unes aux autres. Il convient donc de replacer les initiatives d'évaluation dans un cadre plus général, élargi à plusieurs périodes de programmation et à plusieurs niveaux de responsabilité et alors il est recommandé au sein d'une même évaluation de sélectionner des questions évaluatives dans une perspective ex-ante, mi-parcours ou ex-post. Pour la réussite d'un projet, il est essentiel que des questions évaluatives prioritaires soient précisées dans le mandat donné au groupe de pilotage.

1.1. La valeur ajoutée de l'évaluation

L'évaluation fait désormais partie d'une nouvelle conception de la gestion publique, plus décentralisée et plus responsable. L'évaluation rassemble les informations existantes en un tout cohérent et, le cas échéant, elle en produit de

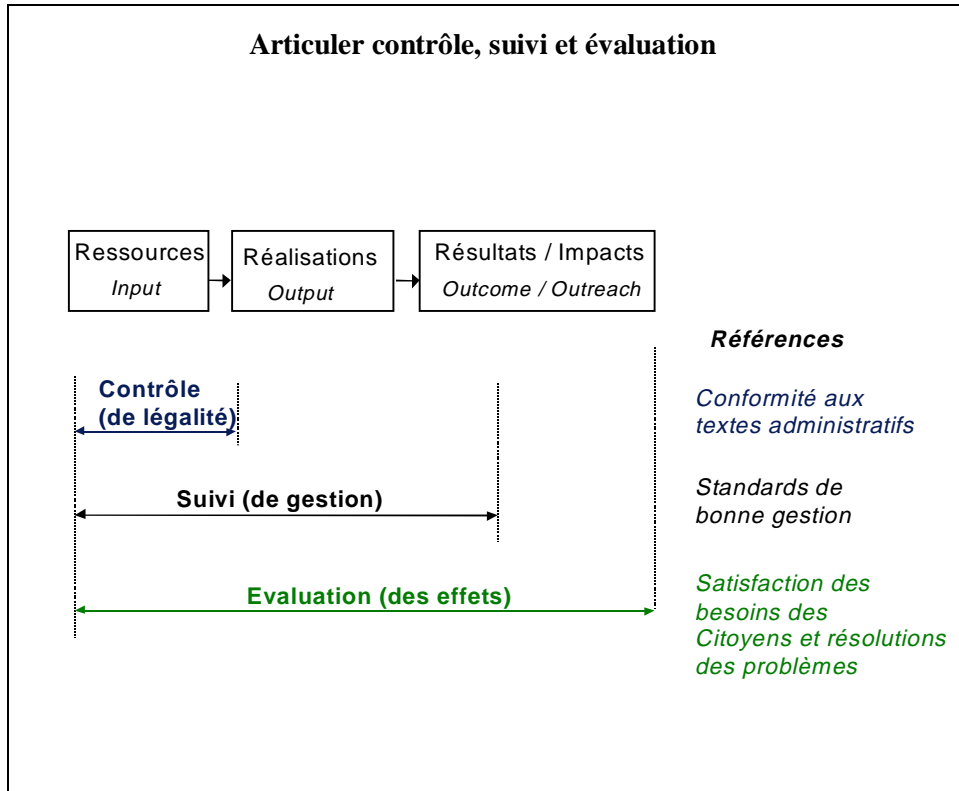
(1) Eric Monnier est directeur du Centre d'expertise en évaluation (C3E), 13 bis, place Jules-Ferry, 69006 Lyon.

nouvelles afin d'accroître la connaissance des responsables sur les réalités sociales et économiques dans lesquelles s'inscrivent les interventions publiques. En outre, l'évaluation est un instrument au service des décisions stratégiques et opérationnelles : son utilité pratique peut se définir comme une mise en relation des objectifs, des moyens, des résultats et des effets des interventions en vue d'améliorer l'efficacité du programme et son efficience. Enfin, l'évaluation introduit la notion de valeur, et *a fortiori* de jugement de valeur, dans les pratiques administratives.

Il faut attendre de l'évaluation ce qu'elle peut réellement apporter et bien préciser sa valeur ajoutée par rapport à des exercices connexes qui ne sont pas les siens, tels que le contrôle et le suivi.

Une première clarification concerne le point de vue à partir duquel l'action publique est jugée. Le contrôle vérifie la légalité et la régularité de la mise en œuvre des ressources. Le suivi quant à lui vérifie la bonne gestion des interventions et produit une analyse régulière sur l'état d'avancement des réalisations. L'évaluation pour sa part sort de la sphère de la mise en œuvre du programme pour juger celui-ci sur la base des résultats{ex "résultat"} et des impacts{ex "impact"} qu'il a produits dans la société.

Un deuxième éclaircissement concerne les critères de jugement. Le contrôle{ex "contrôle"} juge en fonction de critères connus et clarifiés à l'avance (enveloppes budgétaires, textes réglementaires, normes professionnelles). Le suivi juge en fonction d'objectifs opérationnels à réaliser. Au contraire, l'évaluation doit le plus souvent commencer par choisir ses critères de jugement. Ceux-ci sont formulés à partir des objectifs de l'action publique évaluée. Or l'expérience nous montre, et la littérature de l'évaluation le confirme : il est encore rare que les objectifs d'une action publique soient clairs, quantifiés, vérifiables et immédiatement utilisables comme critères d'évaluation. Ces différences fondamentales qui distinguent contrôle, suivi et évaluation sont présentées dans le schéma ci-après.



Les qualités professionnelles et les compétences requises par les différents exercices ne sont donc pas les mêmes. Le contrôleur a une connaissance approfondie des critères préétablis (connaissances juridiques ou comptables par exemple). Son comportement inspire le respect et l'autorité et il n'hésite pas à utiliser la contrainte et à stigmatiser publiquement les erreurs et les dysfonctionnements pour obtenir un effet de dissuasion. Le responsable du suivi doit avoir pour sa part des compétences reconnues en termes d'organisation, de gestion et de mise en place des dispositifs d'information. Il supervise et assiste les opérateurs dans la mesure où il leur fournit des outils et des procédures qui les aident dans leurs pratiques professionnelles et qui de plus serviront à une analyse critique des conditions de mise en œuvre. Enfin, le praticien de l'évaluation ne pouvant que difficilement s'appuyer sur des critères unanimement reconnus, a besoin de la collaboration des gestionnaires publics pour collecter une information pertinente, pour savoir l'interpréter et pour formuler un jugement légitime. Il ne peut pas être efficace si ses interlocuteurs ne sont pas de bonne volonté et de bonne foi, ce qui implique souvent de respecter une certaine confidentialité.

Si les trois exercices présentent des différences intrinsèques, ils se servent mutuellement grâce à leur complémentarité. Par exemple, le contrôle et le suivi peuvent repérer des succès et des échecs apparents, que l'évaluation va chercher à confirmer par une analyse approfondie. À l'inverse, une évaluation peut montrer que l'application d'une norme juridique ou professionnelle constitue un facteur particulier d'efficacité ou d'inefficacité. Entre les exercices de suivi et d'évaluation il existe une interdépendance en ce sens que les indicateurs{ex "indicateur"} du suivi sont utiles à l'évaluation.

Toutefois, les risques de conflits entre les trois logiques sont importants. En particulier, une évaluation ne doit pas être réalisée avec les méthodes du contrôle sous peine de créer des oppositions qui contrarient la collecte d'information et l'utilisation des conclusions. Dans un autre ordre d'idées, aucune évaluation ne peut remplacer un contrôle ou un suivi défaillant. Par exemple, en cas de soupçon sur la régularité ou le professionnalisme de la mise en œuvre, il est nécessaire de procéder à une inspection ou à un audit afin de prendre les décisions qui s'imposent et d'attendre que la sérénité soit revenue pour investir dans une évaluation.

Exemple

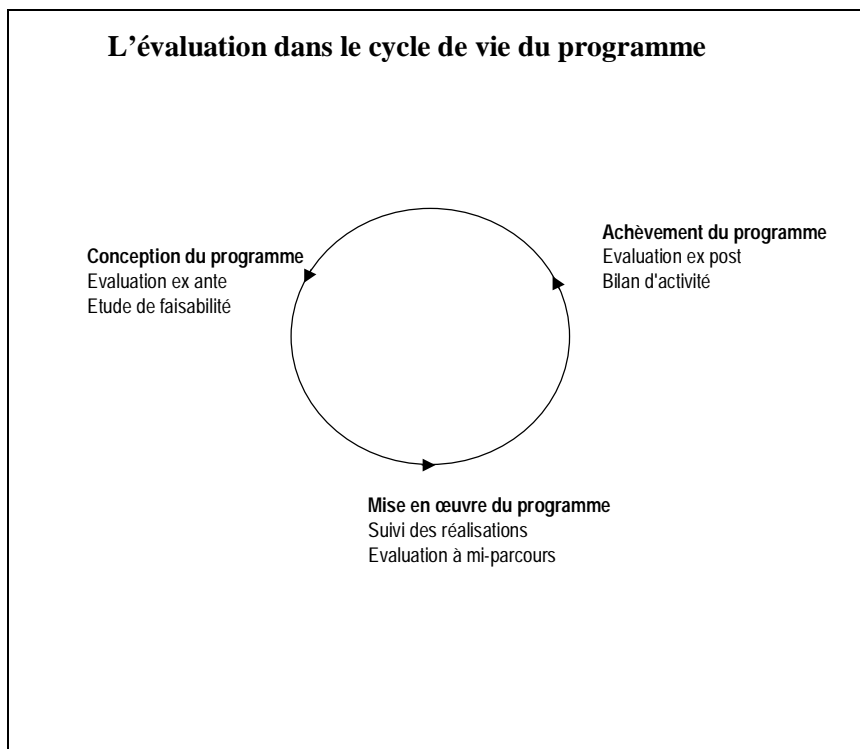
Quelques malentendus à l'égard de l'évaluation intermédiaire d'un programme européen en Bourgogne (France)

Les décideurs interrogés sur leurs attentes à l'égard de l'évaluation ont énuméré, entre autres, les points suivants qui relèvent plutôt du contrôle (C) ou du suivi (S) que de l'évaluation du programme.

- Connaître, de manière synthétique, le contenu des actions, la nature des interventions et les résultats (S).
- Pouvoir rendre compte de l'avancement du programme (S).
- S'assurer que les fonds réservés pour l'agriculture ne sont pas utilisés pour construire des infrastructures (C).
- Mettre à jour les dysfonctionnements du système d'affectation des fonds (C).
- Aider à faire en sorte que l'argent soit dépensé à temps (S ou C).

1.2. Relier les différents types d'évaluation

Pour bien anticiper l'utilité future d'une évaluation, il faut la replacer dans le cycle de vie du programme considéré.



1.3. L'évaluation ex-ante

L'évaluation ex-ante{ex "évaluation ex ante"} intervient en début de cycle alors que le programme n'a pas encore été adopté. Cette forme d'évaluation contribue à faire en sorte que le programme définitif soit aussi pertinent et cohérent que possible. Ses conclusions sont destinées à être intégrées au programme au moment de la décision.

L'évaluation ex-ante porte notamment sur l'analyse des forces, des faiblesses et des potentialités de la région ou du secteur concerné. Elle apporte aux autorités responsables un jugement anticipé sur le fait que les enjeux de développement ont été correctement diagnostiqués, que la stratégie et les objectifs proposés sont pertinents, qu'il n'y a pas d'incohérence avec d'autres politiques, que les impacts attendus sont réalistes, etc. Elle apporte par ailleurs les bases nécessaires au suivi et aux futures évaluations en veillant à l'existence d'objectifs explicites et, dans la mesure du possible, quantifiés. Elle aide à

préciser les critères de sélection des projets à retenir et à s'assurer du respect des priorités. Elle contribue, enfin, à assurer la transparence des décisions en permettant d'expliquer clairement les choix opérés et leurs conséquences attendues.

Les évaluations ex-ante sont réalisées à un moment où les autorités publiques multiplient les discussions et les négociations à propos du futur programme. Elles sont donc soumises à des contraintes fortes : pression sur les délais, formalisation floue de la proposition de programme à évaluer, modifications de cette proposition en cours de travail, exigences de confidentialité, etc. L'équipe d'évaluation doit par conséquent être capable d'intervenir avec souplesse et rapidité, tout en sachant mettre en œuvre des techniques d'analyse des besoins et de simulation des effets socio-économiques.

1.4. L'évaluation à mi-parcours

L'évaluation à mi-parcours{ex "évaluation à mi-parcours"} intervient au deuxième stade du cycle de programmation, pendant la période de mise en œuvre des interventions. Au vu de ses conclusions, des ajustements sont apportés, si nécessaire, en milieu de période. Cette évaluation porte un regard critique sur les premières réalisations et sur les premiers résultats des interventions. Elle apprécie également la qualité du suivi et de la mise en œuvre. Elle montre la traduction opérationnelle des intentions de départ et, le cas échéant, elle constate les modifications apportées de facto aux objectifs initiaux. Par comparaison avec la situation de départ, elle montre l'évolution du contexte économique et social général et juge si les objectifs restent pertinents. Elle examine si l'évolution des politiques et priorités{ex "priorité communautaire"} pose des problèmes de cohérence. Elle aide à préparer des ajustements et des reprogrammations et à les argumenter de façon transparente{ex "transparence"}.

L'évaluation à mi-parcours s'appuie fortement sur les informations issues du système de suivi, mais également sur l'évaluation ex-ante, les informations sur le contexte et son évolution. Ce sont en règle générale des exercices relativement courts et exhaustifs qui s'intéressent surtout aux résultats du programme évalué sans chercher à faire une analyse approfondie de ses impacts. En effet, à mi-parcours d'un programme, les impacts n'ont pas souvent eu le temps de se révéler. Il est toutefois possible et souhaitable de se référer à des évaluations approfondies ou thématiques portant sur les programmes antérieurs lorsque ces analyses existent. L'évaluation à mi-parcours a un caractère « endo-formatif », c'est-à-dire qu'elle alimente une rétroaction directe sur le programme dont elle contribue à améliorer la gestion.

1.5. L'évaluation ex-post

L'évaluation ex-post {ex "évaluation ex post"} récapitule {ex "évaluation récapitulative"} et juge l'ensemble du programme, et notamment ses impacts. Elle vise à rendre compte {ex "rendre compte"} de l'utilisation des ressources, de l'efficacité et de l'efficience des interventions et de l'obtention des effets attendus et d'autres inattendus. Elle porte sur les facteurs de réussite ou d'échec, sur la durabilité des résultats et impacts. Elle cherche à tirer des enseignements généralisables à d'autres programmes ou à d'autres régions.

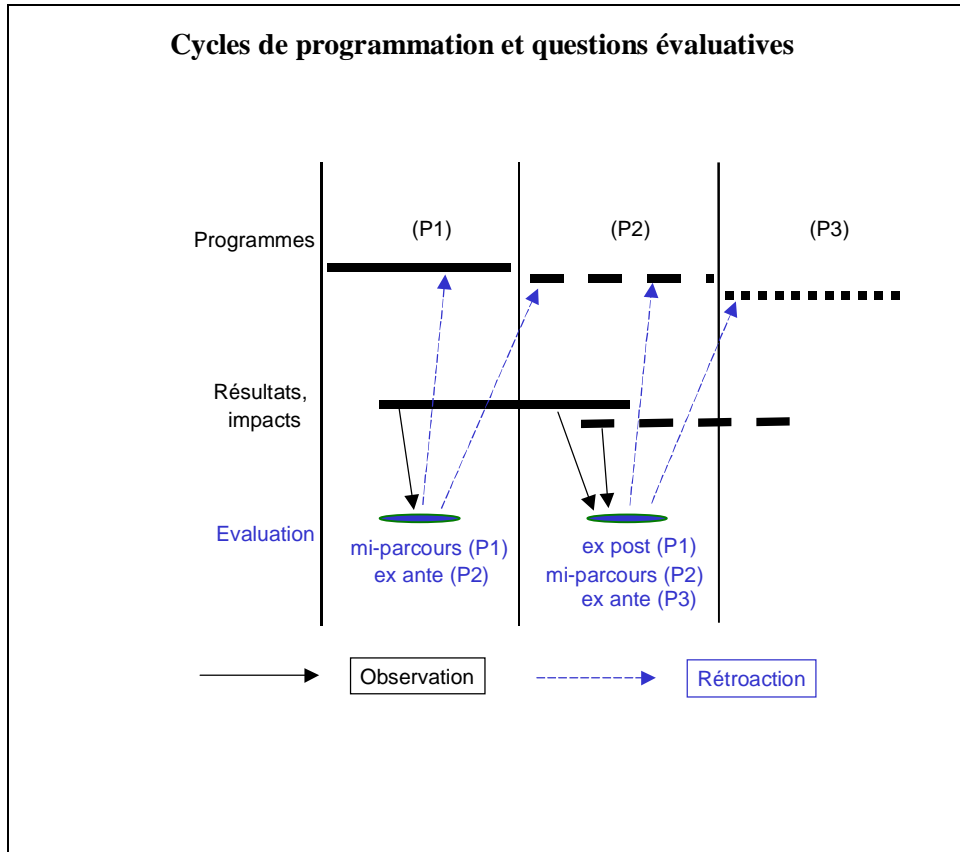
Dans l'idéal, il faudrait disposer de cette évaluation au moment de préparer la prochaine programmation, c'est-à-dire un an au moins avant la date d'achèvement du programme.

Or, pour que les impacts aient eu le temps de se matérialiser, l'évaluation ex-post doit être conduite deux à trois ans après la fin de la période de programmation. À défaut de pouvoir attendre ce délai, il est souvent demandé de réaliser une mise à jour de l'évaluation à mi-parcours, peu de temps avant l'achèvement du cycle de programmation, en liaison avec l'évaluation ex-ante du cycle suivant.

L'analyse des impacts est toujours un exercice de grande envergure s'il est conduit systématiquement. Les évaluations ex-post sont donc susceptibles d'impliquer des enquêtes de terrain et de se dérouler sur des périodes longues allant de douze à dix-huit mois.

2. Les cycles de programmation successifs

L'enchaînement des trois étapes d'évaluation au cours de cycles successifs crée des superpositions qu'il s'agit d'organiser au mieux pour ne pas dupliquer les travaux. Le principe de base revient à adjoindre aux travaux d'évaluation, au cours d'un programme, l'exploitation des conclusions d'évaluations conduites sur le programme précédent. La relative continuité des actions programmées d'une période à l'autre permet, en effet, de tirer des enseignements d'un passé proche sur la pertinence des nouvelles mesures proposées. Le schéma de l'encadré ci-après montre que les interactions sont possibles entre les travaux d'évaluation qui interviennent aux différentes phases de plusieurs programmes successifs.



La mise en œuvre de trois évaluations successives, ex-ante, à mi-parcours et ex-post conduit souvent à une dilution des ressources et à la réalisation de trois analyses relativement superficielles. La notion de l'évaluation selon ces trois stades du cycle de programmation pourrait être remplacée par la notion de questions évaluatives ex-ante pour P3, mi-parcours pour P2 et ex-post pour P1. De cette façon, les informations habituellement rassemblées dans trois évaluations différentes seraient analysées dans un seul projet d'évaluation qui fournirait des éléments sur les trois perspectives temporelles. Dans cet esprit, chaque évaluation serait conçue comme une réflexion sur l'ensemble du programme depuis sa conception jusqu'à la fin de sa réalisation.

3. Formuler les questions évaluatives

3.1. Liste des questions possibles

L'aspect qui est peut-être le plus fondamental dans la phase de conception d'une évaluation est la formulation des questions évaluatives. Ce sont ces questions qui constituent le cœur de l'exercice d'évaluation.

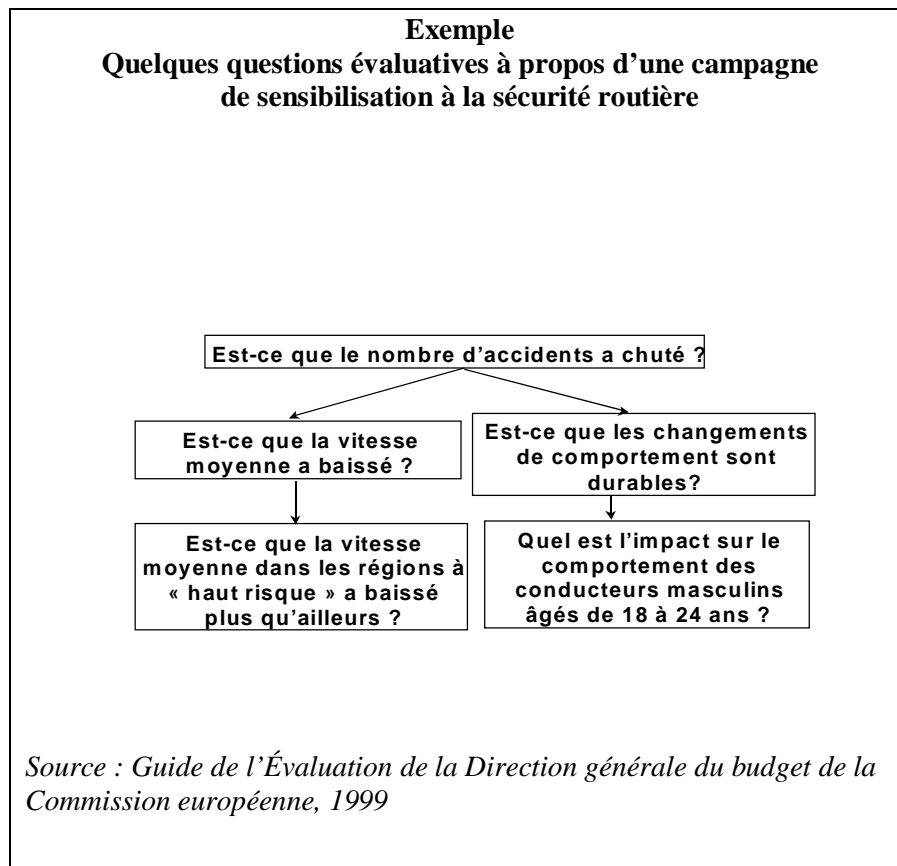
Parmi les questions possibles, on distingue en général les questions descriptives qui cherchent à observer et à mesurer des changements (que s'est-il passé ?), les questions causales qui cherchent à comprendre et apprécier des liens de cause à effet (comment et dans quelle mesure ce qui s'est passé est attribuable au programme ?), et enfin les questions normatives qui appliquent des critères d'évaluation (les résultats et les impacts sont-ils satisfaisants ?). Dans l'idéal, toute question évaluative devrait englober ces trois dimensions et avoir les qualités suivantes :

- La question doit correspondre à une véritable interrogation, à un réel besoin d'information, de compréhension et/ou d'identification de solutions nouvelles. Dans le cas contraire, il s'agit d'une « question alibi », visant par exemple à cautionner une décision déjà prise. Si une question présente un intérêt uniquement en terme de connaissances nouvelles sans débouché immédiat en termes de décisions ou de débat publics, elle relève plus de la recherche scientifique, elle ne trouve pas sa place dans un exercice d'évaluation.
- La question concerne un impact{ex "impact"}, un groupe d'impacts, un résultat{ex "résultat"} ou un besoin{ex "besoin"}, c'est-à-dire qu'elle concerne au moins en partie des éléments extérieurs au programme, à savoir ses destinataires ou son environnement économique et social. Si une question concerne uniquement la gestion interne des ressources et des réalisations, il est probable qu'elle peut être mieux traitée dans le cadre des exercices de suivi{ex "suivi"} ou de contrôle{ex "contrôle"}.
- La question comporte un critère de jugement et un seul{ex "critère de jugement"}. Ce dernier élément pourra paraître parfois difficile à respecter, mais l'expérience montre pourtant qu'il s'agit d'un facteur clé pour l'utilité de l'évaluation. En l'absence de critère de jugement annoncé dès le départ, le rapport final parvient rarement à déboucher sur des conclusions.

3.2. Des questions prioritaires

Parce que chercher à répondre de façon approfondie à toutes les questions serait trop coûteux en temps et en investigations, il est toujours recommandé d'établir une hiérarchie parmi l'ensemble des questions évaluatives {ex "question évaluative"} exprimées.

Pour ce faire, il est nécessaire d'apprécier les enjeux associés à chacune de ces questions (connaissance, négociation, décision, communication) afin de pouvoir établir une hiérarchie entre elles. Les probabilités de réponse aux questions et les probabilités d'utilisation de ces réponses doivent guider le choix qui sera d'autant plus « réussi » que les utilisateurs potentiels auront été consultés et se seront appropriés les questions sélectionnées.



Il est bien entendu qu'il vaut mieux obtenir des réponses légèrement imprécises à des questions importantes que des réponses précises à des questions que personne ne se pose et qui risquent *a fortiori* de ne jamais être écoutées.

On considérera qu'une question est prioritaire si la probabilité de réponse est grande et si la probabilité d'utilisation de la réponse l'est aussi. A cette fin, on se posera les interrogations suivantes :

- Les conclusions seront-elles utilisées ? Par qui ? Pour quoi faire (décider, débattre, rendre compte) ? Quand ?
- Est-il politiquement opportun de mener une telle évaluation à un tel moment ou dans un tel contexte ? Y a-t-il une situation conflictuelle qui risque de compromettre l'exercice ?
- Une étude récente n'a-t-elle pas déjà répondu à la majorité des questions ?
- Les priorités en matière d'évaluation ne sont-elles pas trop instables ?

Si nécessaire, ces points peuvent faire l'objet d'une étude préalable à l'évaluation ou étude d'évaluabilité.

3.3. Critères d'évaluation

Lorsqu'elles comprennent un critère de jugement{ex "critère de jugement"}, les questions d'évaluation se rangent principalement dans l'une des quatre catégories suivantes : celles qui ont trait à la pertinence{ex "pertinence"} du programme, celles qui ont trait à son efficacité{ex "efficacité"}, celles qui ont trait à son efficience{ex "efficience"} ou encore celles qui ont trait à son utilité{ex "utilité"}.

Par « pertinence » on entend, dans le contexte d'une évaluation, l'adéquation entre les objectifs explicites du programme et les problèmes socio-économiques que ce dernier est censé prendre en charge. Dans une perspective ex-ante{ex "évaluation ex ante"}, les questions de pertinence sont les plus importantes car on s'interroge principalement sur la stratégie à choisir ou sur la justification de celle que l'on propose. Dans le cadre d'une analyse à mi-parcours{ex "évaluation à mi-parcours"}, on souhaitera vérifier si le contexte socio-économique a bien évolué comme on l'imaginait a priori et si cette évolution ne remet pas en question tel ou tel objectif initial.

Les questions fondamentales à mi-parcours{ex "évaluation à mi-parcours"} et plus encore ex-post{ex "évaluation ex post"} sont relatives à l'efficacité et à l'efficience des interventions mises en œuvre et du programme dans son entier.

En ce qui concerne l'efficacité, il s'agit de savoir si les objectifs formulés dans le programme sont en voie d'être atteints, de connaître les réussites et les difficultés rencontrées, l'adéquation des solutions choisies ou encore l'influence de facteurs concurrents exogènes au programme lui-même.

L'efficience dans le contexte de l'évaluation revient à rapprocher les résultats obtenus, ou mieux encore les impacts produits des ressources mobilisées. Autrement dit, les effets obtenus sont-ils à la hauteur des moyens engagés ? Les termes d'efficience et d'efficacité sont, par ailleurs, couramment employés par les gestionnaires dans une acception différente dans la mesure où ils s'attacheront, dans le cadre du suivi, à l'analyse des réalisations et non des résultats ou impacts. Soulignons à cet égard que les questions relatives à la performance d'un programme sont de plus en plus fréquentes. Cette notion englobe celles d'efficacité et d'efficience telles que nous les avons définies à la fois pour le suivi et pour l'évaluation. C'est donc un ensemble assez large de questions que l'on trouvera regroupées sous ce vocable.

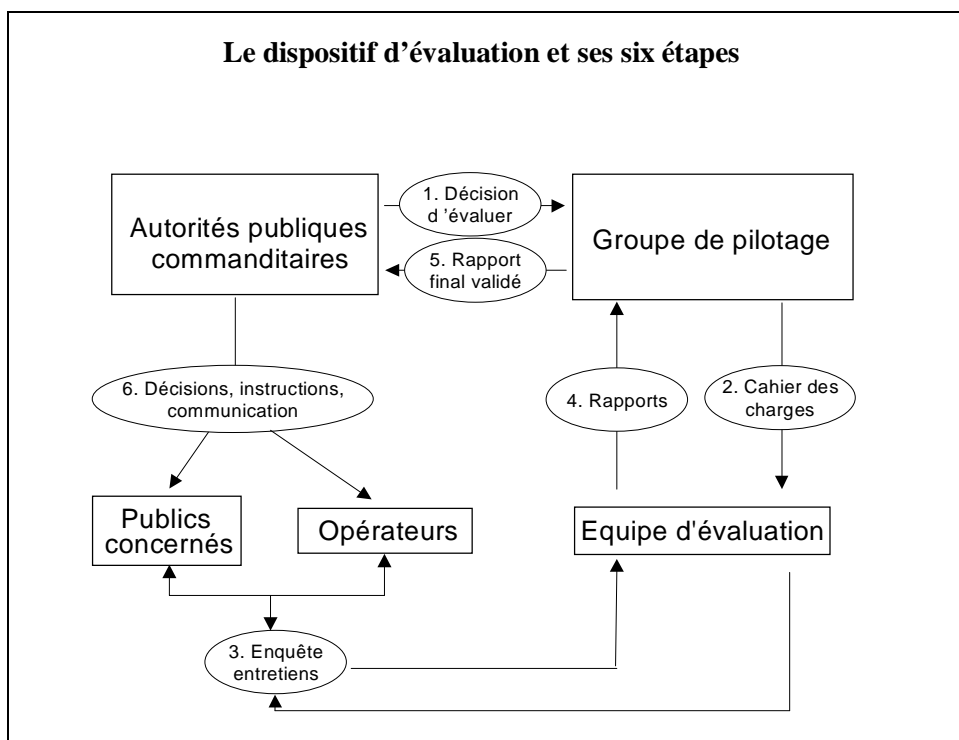
Le critère d'utilité juge les impacts obtenus par le programme au regard des besoins de la société et des enjeux économiques. L'utilité est un critère d'évaluation très particulier puisqu'il abandonne toute référence aux objectifs officiels du programme. Il peut être judicieux de formuler une question d'utilité lorsque les objectifs sont mal définis ou lorsque qu'il y a beaucoup d'impacts imprévus. Ce critère doit malgré tout être utilisé avec précaution afin d'éviter que l'équipe d'évaluation ne procède à une sélection trop personnelle des besoins ou enjeux importants. Certains auteurs nord-américains ont prôné cette forme d'évaluation « affranchie des objectifs » (*goal free evaluation*).

4. Répartir les rôles dans le dispositif d'évaluation

On entend par « dispositif d'évaluation » {ex "dispositif d'évaluation"} les rôles et responsabilités au sein du processus d'évaluation d'une part et, d'autre part, les relations entre l'évaluation et la gestion du programme. Des règles du jeu claires et équitables contribuent à des travaux de qualité et évitent les blocages. Le choix principal porte sur le rôle de pilotage de l'évaluation, rôle qui comprend notamment la rédaction du cahier des charges, la sélection de l'équipe d'évaluation et la validation de son rapport.

Par le passé, le pilotage de l'évaluation a été souvent confié à un fonctionnaire isolé ou au contraire, exercé collectivement par les cofinanceurs du programme dans le cadre de réunions formelles, élargies à de nombreux partenaires et peu fréquentes {ex "comité de suivi"}. En pratique, le dispositif fonctionne bien

quand la responsabilité du pilotage est assurée par un groupe restreint{ex "groupe de pilotage"} d'une dizaine de fonctionnaires et de porte-paroles des publics les plus directement concernés par le programme, auquel on adjoint parfois quelques experts. Les réunions doivent être fréquentes (une par mois à certains moments) et les procédures de travail souples et peu formelles.



De façon schématique (voir schéma ci-dessus) on peut résumer le dispositif d'évaluation{ex "dispositif d'évaluation"} à la répartition des rôles entre les cinq entités clés que sont les commanditaires, le groupe de pilotage, l'équipe d'évaluation et enfin les opérateurs et les publics concernés. Dans certaines situations il convient de rajouter d'autres entités. En particulier on a parfois, avec succès, organisé le pilotage sur la base d'un groupe global pour l'évaluation du programme coordonnant plusieurs groupes sectoriels (ou des comités créés à une échelle géographique plus petite). Les responsabilités des entités de base sont décrites par leur rôle respectif au cours des six étapes clés qui constituent une évaluation pluraliste¹.

¹ Cf. pour plus de développement sur la démarche pluraliste : Monnier, (E.), « Évaluation de l'action des pouvoirs publics », *Economica*, 1992.

- Étape 1 : Les autorités publiques, commanditaires et futurs utilisateurs de l'évaluation, formulent leurs instructions à un groupe de pilotage{ex "groupe de pilotage"}. Cette étape se matérialise a minima par un paragraphe dans le document de programmation et si possible par un mandat au groupe de pilotage.
- Étape 2 : Le cahier des charges{ex "cahier des charges"} est conçu et rédigé par le groupe de pilotage{ex "groupe de pilotage"} sur la base du mandat reçu. Ce cahier des charges sert le plus souvent de document d'appel d'offres et permet la sélection de l'équipe d'évaluation. Le groupe de pilotage assurera autant d'interactions que nécessaires avec l'équipe d'évaluation afin de s'assurer que celle-ci ne dévie pas des lignes directrices inscrites au cahier des charges et travaille dans le sens des questions posées.
- Étape 3 : L'équipe d'évaluation réalise la collecte d'informations secondaires et primaires auprès des opérateurs et des publics concernés puis en fait l'analyse. C'est l'étape la plus longue.
- Étape 4 : Un rapport que l'on peut qualifier de « technique » est rédigé par l'Équipe d'évaluation. Il inclut des conclusions et le cas échéant des recommandations provisoires à l'intention du groupe de pilotage{ex "groupe de pilotage"}.
- Étape 5 : Les conclusions sont validées en tout ou partie par le groupe de pilotage{ex "groupe de pilotage"} qui en outre formule ses recommandations aux autorités publiques dans le rapport d'évaluation qui intègre le rapport technique ou le présente en annexe.
- Étape 6 : Les autorités utilisent l'évaluation en la diffusant auprès du public et/ou en décidant de nouvelles orientations que les administrations seront chargées de mettre en œuvre. Cette étape cruciale est souvent isolée du processus d'évaluation proprement dit car le groupe de pilotage peut être dissous après avoir rendu son rapport et ne peut malheureusement pas assurer un véritable suivi (*follow-up*) des décisions subséquentes. Notons toutefois que l'on peut se prémunir contre le risque d'absence de suites données en incluant dans le mandat initial un engagement moral des autorités d'inviter si ce n'est le groupe de pilotage au moins son président à présenter les conclusions et à assister à la réunion au cours de laquelle seront prises les décisions.

Il faut souligner que, par le passé, les dispositifs d'évaluation ont peu à peu convergé vers le schéma qui précède pour les évaluations à mi-parcours. Par contre tel n'a pas été le cas des évaluations ex-ante{ex "évaluation ex ante"}.

Les contraintes de temps et la difficulté de mobiliser les protagonistes par rapport à un programme qui n'existe pas encore peuvent expliquer sans toutefois complètement justifier l'absence de groupe de pilotage.

Les mêmes lacunes se retrouvent en matière d'évaluation ex-post{ex "évaluation ex post"} alors même que l'on ne peut se prévaloir de difficultés similaires. Ce type d'évaluation est souvent demandé par une administration sans véritablement apporter tout le soin nécessaire à la clarification des attentes des niveaux décisionnels les plus importants ou au pilotage du processus.

5. Vers un renforcement de la dimension politique ?

Les règlements européens ont entraînés une dynamique spécifique concernant la gestion et l'évaluation des programmes. Les services de la Commission ont fourni des efforts remarquables en matière d'évaluation, par eux-mêmes mais aussi en étroite coopération avec les administrations nationales. Des progrès notables ont été réalisés pendant ces dernières années : le développement quantitatif des évaluations s'est accompagné d'un développement qualitatif et plus de professionnalisme.

Le partenariat avec les services de la Commission a joué un rôle positif en créant des relations de confiance. Il a crédibilisé l'évaluation et en a ainsi facilité la mise en œuvre.

On a constaté toutefois que les spécificités institutionnelles de chaque État conduisent à des différences importantes du point de vue de la mise en œuvre de l'évaluation. De manière un peu schématique, on peut discerner trois phases dans le niveau d'utilité attribué à l'évaluation et par là même, dans la qualité de son organisation.

1^{re} phase : L'évaluation est vécue comme une réponse à des obligations réglementaires. Elle est perçue comme une contrainte et une charge de travail supplémentaire qui pèsent sur les gestionnaires.

2^e phase : L'évaluation devient un système d'aide à la conception et à la gestion des interventions. Un dialogue s'établit avec les évaluateurs et la qualité des informations recueillies s'améliore.

3^e phase : L'évaluation devient un acte politique et ses résultats font l'objet de débats publics. Le but recherché est d'informer l'opinion publique sur

l'efficacité de l'utilisation des fonds publics et sur la démonstration de leur efficacité en termes d'obtention des impacts attendus (*Value For Money*).

Il est intéressant de remarquer que l'utilisation démocratique{ex "démocratique"} de l'évaluation est favorisée lorsqu'il existe une pré-culture indépendamment des exigences communautaires. Autrement dit, l'obligation réglementaire peut aider à établir une utilisation managériale mais n'est probablement pas suffisante pour induire une contribution au débat public.

6. Panorama des pratiques d'évaluation en Europe par Jacques Toulemonde¹

Cet article reprend le contenu d'une présentation faite successivement au Conseil national de l'évaluation le 25/11/1999 et à l'École nationale d'administration le 23/02/2000. La préparation de ces deux conférences était une occasion d'organiser mes idées à propos du développement des pratiques d'évaluation en Europe.

Mes idées ont été inspirées par une série de travaux auxquels j'ai participé depuis près de dix ans. Il s'agit des travaux suivants :

- Les deux enquêtes réalisées dans le cadre du programme MEANS de la Direction générale politique régionale de la Commission européenne (Commission européenne, 1999). Ces enquêtes ont eu lieu respectivement en 1992 et 1997. La seconde a couvert les quinze États membres, avec plus de cent entretiens téléphoniques et des visites dans la plupart des capitales européennes. Elle a couvert le champ des Fonds structurels, qui s'étend à une part significative des politiques publiques dans la plupart des États membres.
- La participation de notre équipe aux activités de plusieurs sociétés professionnelles (Société européenne d'évaluation, société britannique, suisse, italienne et plus récemment française et wallonne), ainsi que ma participation à l'intéressant séminaire franco-britannique organisé par le Conseil scientifique de l'évaluation en janvier 1998 (CSE, 1999).
- Diverses missions d'assistance réalisées par notre Centre pour la Commission européenne et comportant une appréciation sur la qualité de travaux réalisés par des équipes de différentes nationalités.

(1) J. Toulemonde est directeur associé au Centre européen d'expertise en évaluation, 13 bis, Place Jules-Ferry, 69006 Lyon.

- Ma participation à la préparation d'un ouvrage collectif intitulé « International Evaluation Atlas », qui doit paraître prochainement (Furubo, Rist, Sandhal, à paraître). Cet ouvrage, réalisé au sein de l'International Evaluation Research Group (INTEVAL), présente en 23 chapitres l'état de la pratique d'évaluation dans les principales démocraties et institutions internationales. Il couvre l'Union européenne, les principaux pays qui la composent, ainsi que la Suisse et la Norvège.

La conception de cet atlas nous a posé un problème qui se répète dans les mêmes termes à l'occasion de cette présentation : y a-t-il des spécificités nationales dans les pratiques d'évaluation, et si oui, celles-ci prévalent-elles sur les spécificités sectorielles ? Cette question est incontournable dans la mesure où les pratiques d'évaluation varient très fortement d'un domaine politique à l'autre. On n'évalue pas une politique de vaccination comme un projet de recherche ou un programme de reconversion industrielle. Peut-on raisonnablement rechercher des constantes nationales dans ce patchwork de pratiques hétéroclites ?

Il sera répondu à cette question plus loin. Contentons-nous pour l'instant de prendre pour hypothèse de travail que les spécificités nationales sont à la fois identifiables et suffisamment significatives pour donner lieu à des comparaisons intéressantes.

La présentation qui suit adopte successivement deux angles de vue :

- Les facteurs du développement de l'évaluation : comment l'évaluation est-elle née et comment a-t-elle évolué dans chaque pays ?
- L'acculturation à l'évaluation : jusqu'à quel point l'évaluation a-t-elle changé les règles du jeu politico-administratif dans les différents pays ?

Ces deux points de vue portent sur un seul et même paysage de l'évaluation européenne. Les descriptions qui suivent se recouvrent donc en partie. Espérons que cela précisera le relief sans alourdir la lecture.

6.1. L'évaluation, produit d'importation ou initiative nationale ?

L'évaluation importée des États-Unis dans les années 1970

Chacun sait que l'évaluation est née aux États-Unis dans le même berceau que le Planning-Programming-Budgeting-System (PPBS). Elle a donc été importée dans les années 1970, au moment où le PPBS a été francisé sous la dénomination de rationalisation des choix budgétaires. À la même époque, la

plupart des pays de l'Europe du Nord ont fait de même en instituant des unités ou des commissions chargées de l'analyse des politiques publiques. Ces unités avaient généralement une mission d'évaluation ex-ante et ex-post, avec une vocation interministérielle et une ambition affichée de faire prévaloir plus de rationalité scientifique dans les arbitrages budgétaires. Des pays comme le Royaume-Uni ou les Pays-Bas ont montré le même empressement que la France à décalquer ce modèle américain, à un moment où Wildavsky (1969) avait déjà prédit son caractère inapplicable !

Tous les pays qui ont connu cette forme d'évaluation l'ont maintenant dépassée, sinon abandonnée. C'est en France que la transition s'est faite de la façon la plus tranchée. En abandonnant la RCB au début des années 1980, on a jeté le bébé avec l'eau du bain, au point que l'évaluation a perdu tout soutien significatif à l'échelon gouvernemental pendant près de dix ans. Dans d'autres pays, la transition s'est faite de façon beaucoup plus progressive.

Aux Pays-Bas, par exemple, le ministère des Finances est passé sans rupture d'une pratique fondée sur le PPBS à une pratique beaucoup plus pragmatique qui sera évoquée plus loin. Au Royaume-Uni, la Commission d'analyse de politique mise en place dans les années 1970, a été stoppée en même temps que le gouvernement de Margaret Thatcher relançait l'évaluation avec le slogan « en avoir pour son argent » (*value for money*).

De façon étonnante, l'Italie a connu la même évolution que la France avec dix ans de décalage. Au début des années 1980, l'évaluation ex-ante y a connu un formidable développement sous la pression d'une obligation gouvernementale qui consistait à faire passer le test de l'analyse coûts-avantages à tous les projets d'investissement publics. L'importation s'est faite en provenance de la Banque mondiale à un moment où l'administration publique était particulièrement discréditée. Ici aussi, la rationalité technique et économique n'a pas résisté aux forces politiques. Au pays de Machiavel, il n'a pas fallu dix ans pour que se montent des cabinets capables de réaliser une analyse coût-avantages favorable à n'importe quel projet. La pratique a tourné au rituel et a été abandonnée.

Le développement au sein des réseaux professionnels internationaux

À la fin des années 1970, la culture de l'évaluation s'est répandue dans quelques milieux professionnels particulièrement ouverts aux courants internationaux, en particulier ceux de la recherche et de l'aide aux pays en développement.

C'est dans ces deux domaines que la plupart des pays européens, ainsi que la Commission elle-même, ont commencé à construire une véritable capacité

d'évaluation. Ainsi, pendant près de quinze ans, la seule véritable expertise italienne en matière d'évaluation de politique et de programme était cantonnée aux politiques de recherche. Aujourd'hui encore en Allemagne et au niveau fédéral, ce sont les politiques scientifiques et l'aide au développement qui font l'objet des systèmes d'évaluation les plus cohérents.

Que ce soit pour la recherche ou l'aide au développement, l'OCDE a joué un rôle clé dans la diffusion de l'évaluation en réunissant des groupes de travail internationaux et en publiant des ouvrages de référence. Par exemple, le Comité d'aide au développement de l'OCDE a un groupe de travail consacré à l'évaluation depuis de nombreuses années.

Paradoxalement, cette large diffusion de l'évaluation à l'échelle européenne n'a pas été un véritable moteur de développement des pratiques. Les réseaux professionnels sont en effet restés très compartimentés et peu enclins à jeter des ponts vers les autres secteurs. Chaque domaine politique a développé ses propres outils (panels d'experts pour l'évaluation de la recherche, cadre logique pour l'évaluation de l'aide au développement). Chaque domaine a généré son réseau d'experts et de spécialistes sans contribuer de façon décisive à généraliser la pratique de l'évaluation.

L'évaluation imposée par les Fonds structurels européens

Il n'en a pas été de même de la politique régionale de l'Union européenne, mise en œuvre au moyen des Fonds structurels (FEDER, FSE, etc.). Cette politique a véritablement fait décoller l'évaluation en Europe. C'est pourquoi elle doit être décrite plus en détail.

Depuis le traité de Rome, le niveau européen a mission de coordonner les aides aux régions souffrant d'un retard de développement ou de problèmes de reconversion. Les dépenses budgétaires relatives à cette politique ont pris, et continuent de prendre, une part croissante du budget européen. Dans ce contexte de dépenses publiques de plus en plus lourdes, l'Acte unique européen, adopté au début 1986, a introduit une première référence forte à l'amélioration de l'efficacité des dépenses publiques.

Dans cette ligne, les règlements de 1988 relatifs à la période 1989-1993 des Fonds structurels ont créé une obligation d'évaluation systématique. Ils stipulaient que « en vue d'apprécier l'efficacité des interventions structurelles, l'action communautaire fait l'objet d'une évaluation ex-ante et ex-post destinée à apprécier son impact ».

L'obligation d'évaluer ne s'est pas immédiatement traduite dans la pratique pour plusieurs raisons. D'abord, la gestion des Fonds structurels a pris la forme de programmes pluriannuels et partenariaux, ce qui a constitué une sorte de révolution pour de nombreuses administrations nationales ou régionales. Quant à la culture d'évaluation, elle était étrangère à la plupart des administrations concernées, ainsi d'ailleurs que pour les services de la Commission eux-mêmes, au moins en ce qui concerne les Fonds structurels. Par conséquent, les évaluations réalisées pendant la période 1989-1993 n'ont pas été systématiques, ni de très grande qualité malgré les efforts déployés au niveau de la Commission.

L'impératif de cohésion économique et social a été réaffirmé dans le traité de Maastricht, et un nouveau renforcement de l'effort budgétaire s'est accompagné d'un second appel solennel à garantir l'efficacité des Fonds. À l'instigation de la Commission, mais également sous la forte impulsion de certains États membres (Pays-Bas et Royaume-Uni en particulier), les règlements adoptés en juillet 1993 pour la période 1994-1999 ont renforcé considérablement les exigences en matière d'évaluation. La réglementation a créé pour les autorités nationales et régionales, une obligation d'appréciation ex-ante, de suivi et d'évaluation ex-post.

L'article 26 du règlement de coordination a été jusqu'à établir une clause spécifiant « que les aides seront octroyées lorsque l'appréciation ex-ante aura démontré les avantages socio-économiques à retirer à moyen terme, eu égard aux ressources mobilisées ». Cette exigence très forte de justification économique a priori, variante communautaire de la notion britannique de *Value For Money*, n'est pas sans rappeler la rationalisation des choix budgétaires. Elle est très loin d'avoir fait la preuve de son applicabilité pratique.

Au contraire, les règlements communautaires ont eu une influence considérable sur le développement de l'évaluation, mais sous une forme plus souple. Pour comprendre cette influence, il faut se rappeler que l'obligation d'évaluer s'est imposée à plusieurs centaines d'autorités régionales. Sur le plan quantitatif, la pratique d'évaluation s'est développée de façon spectaculaire. Ainsi, dans les quinze États membres, pour la période de 1996 au début de 1998, plus de 380 évaluations à mi-parcours ont été organisées. Le nombre d'évaluations de programme dans le cadre des Fonds structurels a été multiplié par 5 ou 6 par comparaison avec la période 1989-1993.

En outre, les programmes de développement régional étaient et restent dirigés en priorité vers les pays d'Europe du Sud qui sont les moins familiers avec

l'évaluation. De plus, le spectre ouvert au financement des Fonds structurels étant extrêmement large, leur gestion mobilise la majorité des services administratifs dans les pays et régions concernés.

Avec le renforcement de la réglementation européenne en 1993, l'évaluation a été imposée et reconnue comme un point de passage incontournable pour bénéficier de la politique communautaire. L'évaluation a alors connu un développement considérable, notamment sous la forme d'évaluation intermédiaire (ou à mi-parcours). On peut dire qu'à la fin des années 1990, plus aucun pays européen n'ignore la pratique de l'évaluation, et que la qualité de cette pratique a augmenté de façon significative.

On ne peut cependant que s'interroger sur la stratégie d'obligation systématique qui a été adoptée. Dans le cadre des Fonds structurels, l'obligation s'étend en principe à toutes les interventions et à toutes les étapes du cycle de programmation (ex-ante, intermédiaire, ex-post, sans compter l'ajout récent d'une étape d'évaluation « finale »). Il me semble pouvoir affirmer qu'une obligation aussi astreignante à une aussi grande échelle n'a pas d'équivalent au monde. À l'actif de cette stratégie, il faut indéniablement mettre la très large diffusion actuelle des pratiques d'évaluation en Europe. Le passif consiste en ce que beaucoup de travaux d'évaluation restent encore trop formels et ne comportent pas le minimum d'observation empirique et d'analyse des impacts qui permettraient de juger de l'efficacité réelle des dépenses.

La Banque mondiale et les pays européens en développement

Lors de notre enquête de 1992, nous avons été surpris de trouver de bons connaisseurs de l'évaluation en Grèce et au Portugal, alors même que la culture administrative nationale était très juridique. Il s'est avéré qu'avant leur adhésion à la Communauté, ces pays avaient bénéficié de prêts de la Banque mondiale et d'autres institutions internationales. Comme toutes les aides aux pays en développement, ces financements étaient systématiquement associés à des évaluations ex-ante et ex-post. Bien que très centralisés au niveau des sièges des institutions internationales, ces systèmes d'évaluation n'en avaient pas moins commencé à créer des compétences locales. Le même phénomène a joué et continue de jouer dans les pays d'Europe de l'Est.

Il s'agit là du quatrième et dernier facteur exogène de développement des pratiques d'évaluation. Les trois facteurs qui suivent sont plus endogènes. Ils correspondent à des développements initiés depuis l'intérieur du pays, par des champions nationaux, et non plus importés de l'étranger.

Le rôle inégal des Parlements nationaux

Alors que le Congrès a joué un rôle essentiel dans le développement de l'évaluation aux États-Unis, rares sont les Parlements nationaux qui ont fait de même en Europe. Les parlementaires ont souvent considéré que le jugement sur les interventions publiques appartenait à eux et à eux seuls. La conséquence est que les exercices d'évaluation qu'ils ont conduit ont souvent pris un tour de commissions d'enquête, avec un caractère public et contradictoire, mais une dimension empirique limitée et des critères de jugement peu explicites.

Pratiquement aucun Parlement européen ne s'est équipé d'une capacité d'évaluation propre à un niveau qui lui permettrait de rivaliser avec l'exécutif. À ma connaissance, des services d'évaluation ont été mis en place à une échelle limitée par les Parlements suisses et français, ainsi qu'au Parlement européen. À ce jour ces expériences n'ont pas eu un rôle moteur dans le développement des pratiques.

À l'inverse, le Parlement a eu une influence déterminante en Allemagne et dans les pays scandinaves. Dans le cas de l'Allemagne, les parlementaires ont exigé et obtenu qu'un certain nombre de politiques majeures fassent l'objet de rapports périodiques de la part d'institutions indépendantes de l'exécutif. Cette pratique ne porte pas le nom d'évaluation mais elle en partage souvent la substance : examen empirique des effets, analyses approfondies, jugements sur l'efficacité, l'efficience ou la pertinence. En Scandinavie, les commissions parlementaires exercent également un pouvoir important en imposant la réalisation d'évaluations par des équipes universitaires ou para-publiques indépendantes de l'exécutif.

Les initiatives de quelques Cours des comptes

La plupart des Cours des comptes européennes s'intéressent à l'évaluation et beaucoup de leurs membres y consacrent une part de leur temps. Par contre, très peu de ces institutions ont mis en place une véritable capacité d'évaluation avec une équipe spécialisée, des recrutements adaptés et une capitalisation collective de l'expérience.

À ma connaissance, les exceptions sont le Royaume-Uni, les Pays-Bas et la Suède, encore que dans ce dernier pays, la fonction d'évaluation vienne d'être séparée du contrôle et déplacée vers une agence indépendante.

C'est dans le cas des Pays-Bas que la Cour des comptes me semble avoir eu le plus grand effet d'entraînement. Dans les années 1980, cette institution a créé en son sein une équipe spécialisée qui a réalisé un travail méthodologique

approfondi, ainsi que des méta-évaluations et des appréciations de qualité sur les travaux d'évaluation des différentes administrations. Il s'agit d'une contribution tout à fait significative au développement de la pratique d'évaluation dans le pays. L'apport de la Cour des comptes néerlandaise s'est d'ailleurs étendue au-delà des frontières puisqu'elle a été à l'origine de la Société européenne d'évaluation.

La pression des ministères des Finances

Dans de nombreux pays européens, le ministère des Finances a joué un rôle de champion national de l'évaluation. L'exemple qui vient immédiatement à l'esprit est le Royaume-Uni, où les études de type « value for money » ont été cadrées et poussées dès les années 1980 par le Treasury. Beaucoup plus récemment, le Ministero del Bilancio a joué un rôle décisif dans le développement de l'évaluation en Italie, sous l'impulsion de M. Ciampi, aujourd'hui président de la République. Au niveau de la Commission européenne, les responsables du budget ont également joué le rôle décisif. C'est en effet Erkki Liikanen, Commissaire au Budget, qui a lancé l'initiative SEM 2000 (*Sound and Efficient Management*), initiative qui est à l'origine de la généralisation de l'évaluation au-delà des services précurseurs (aide au développement, recherche et Fonds structurels).

On conçoit aisément que l'évaluation se développe à l'interface entre les ministères dépensiers et le ministère des Finances. Le moteur est alors celui de la justification de la dépense publique. C'est celui qui a entraîné un développement rapide et généralisé de l'évaluation en Australie dans les années 1980. On peut se demander pourquoi le ministère des Finances français ne joue aucun rôle de ce type.

Il est vrai que l'évaluation imposée par les grands argentiers risque de faire retour à la rationalisation des choix budgétaires, c'est-à-dire à l'utopie d'une rationalité économique primant sur la logique politique. Pour éviter ce risque, le ministère des Finances néerlandais a mis en place, dès la fin du PPBS, un système souple et original appelé « procédure de reconsidération ». Tous les cinq ans, chaque ministère dépensier est tenu de réaliser des évaluations ex-post incluant des recommandations sur les programmes ou mesures qu'il faudrait éliminer ou réformer dans l'hypothèse d'une réduction budgétaire de 20 %. Il semble que l'enjeu politique reste placé trop haut car, après plusieurs cycles successifs de reconsidérations, un bilan montre qu'une trop petite proportion des évaluations ont été suivies d'effets budgétaires (Toulemonde, 1999).

Première conclusion : les différences nationales tendent à s'estomper

À l'issue de ce premier tour d'horizon, on constate que les pays européens ont subi des influences multiples et souvent convergentes : influence du PPBS, influence de la pratique d'évaluation des projets de développement (pour les pays receveurs aussi bien que pour les pays donateurs), influence des Fonds structurels européens.

À ces facteurs exogènes sont venu s'ajouter selon les pays des initiatives spécifiques : pression du Parlement en Allemagne, initiatives des Cours des comptes aux Pays-Bas, en Suède et plus récemment au Royaume-Uni, pression des ministères des Finances, comme c'est le cas à présent en Italie.

Il est surprenant de constater à quel point les origines et les ambitions de ces « champions nationaux » sont différentes. Quoi de comparable entre l'initiative de gestion financière lancée dans les années 1980 par le gouvernement de Margaret Thatcher et le système interministériel mis en place quelques années plus tard par Michel Rocard ? La première formule visait à gérer au plus juste les deniers des contribuables tandis que la seconde avait pour ambition initiale d'élever la qualité du débat public. Pourtant, les deux formules ont contribué à diffuser la culture d'évaluation dans les administrations de part et d'autre de la Manche.

Ont-elles laissé des traces indélébiles dans les cultures politico-administratives nationales ? Il est permis d'en douter au vu de la multitude des formes d'évaluation qui se pratiquent aujourd'hui dans les deux pays. Même s'il y a eu un style britannique et un style français d'évaluation, les différences ont tendance à s'estomper à mesure que l'activité se professionnalise. Plus l'expérience s'accumule, plus on en tire les leçons, et plus on s'aperçoit que ces leçons sont universelles.

Mon opinion est qu'il existe des spécificités nationales fortes dans la façon dont l'évaluation s'est développée, spécificités qui tiennent au substrat administratif de départ et à des contingences historiques. Par contre, je doute que ces spécificités subsistent lorsque la pratique d'évaluation arrive à maturité.

6.2. L'évaluation : exercice administratif, outil de gestion ou devoir démocratique ?

Cette seconde partie est très directement inspirée de nos enquêtes sur la diffusion de l'évaluation dans des Fonds structurels. On en trouvera la relation originale dans le volume 1 de la Collection MEANS (Commission européenne,

1999). Ces enquêtes ont été une occasion de situer le degré d'acculturation à l'évaluation dans les différents pays européens.

Dans le cadre de la mise en œuvre des Fonds structurels, les pratiques d'évaluation ont rapidement convergé vers le système d'évaluation imposé depuis Bruxelles. Ce système partage les mêmes fondements que l'évaluation des projets d'aide au développement. L'évaluation est conçue comme un jugement indépendant sur la pertinence, l'efficacité et l'efficience des programmes, par référence à des objectifs clarifiés, hiérarchisés et si possible quantifiés.

En s'imposant à tous les États membres, l'obligation d'évaluer a cependant rencontré des contextes et des terrains plus ou moins favorables. Les nouveaux venus sur le champ de l'évaluation, qu'ils soient administrateurs, universitaires ou consultants, ont naturellement fait référence aux schémas de pensée qui leur étaient familiers. On a vu ainsi se dessiner des tendances nationales à concevoir l'évaluation comme un exercice de « *Value For Money* » (Royaume-Uni), comme une analyse stratégique (France) ou comme une analyse économique de projet (Italie).

La différence la plus importante ne tient cependant pas à la manière de conduire l'évaluation, mais à celle de l'utiliser. De façon un peu schématique, on peut discerner trois phases dans le niveau d'appropriation de l'évaluation par les acteurs politico-administratifs.

6.3. L'évaluation vécue comme une obligation administrative

Même si l'obligation d'évaluer résulte de décisions adoptées à l'unanimité des États membres, elle a d'abord été vécue dans certains pays comme une contrainte. Nombre de gestionnaires publics l'ont considérée comme une charge de travail supplémentaire exigée par la Commission.

En s'en tenant à l'image instantanée de notre enquête réalisée en 1997-1998, on doit admettre que cet état d'esprit se rencontre dans nombre de régions d'Europe du Sud. Il se traduit par des pratiques du style de cette région italienne qui a commandé à un consultant l'évaluation d'un programme de plus d'un milliard d'euros, en lui demandant de finir le travail en quelques semaines. Cette tactique a été spontanément adoptée par ceux qui souhaitent des évaluations sans danger, interférant le moins possible avec la gestion administrative et les décisions politiques.

Il n'est cependant pas juste de présenter une telle image instantanée sans évoquer les dynamiques propres à chaque pays. Ainsi, l'administration portugaise s'est considérablement familiarisée avec l'exercice d'évaluation, et au cours des dernières années, d'importants progrès ont été réalisés en termes de qualité des travaux et d'utilisation des conclusions. De même, le gouvernement italien vient d'ajouter sa propre pression à celle de la Commission pour que les autorités régionales prennent l'évaluation au sérieux.

L'évaluation comme outil d'aide à la gestion publique

Dans la majorité des régions bénéficiaires des Fonds structurels, les évaluations réalisées en 1996-1997 ont suscité l'intérêt des autorités gestionnaires des programmes et provoqué l'émergence d'une nouvelle activité pour les consultants et les centres de recherche universitaires.

Dans de très nombreux cas, les responsables des programmes ont perçu l'équipe d'évaluation externe, non pas comme une strate de bureaucratie supplémentaire, mais comme une source de conseil pour leur gestion et comme une médiation entre les partenaires publics concernés. Des relations de confiance ont été établies, la qualité des informations recueillies s'est renforcée et l'utilisation des conclusions s'est renforcée.

De nombreux gestionnaires régionaux sont allés au-delà des exigences réglementaires en approfondissant leurs travaux d'évaluation dans tel ou tel domaine. Les administrations nationales ont pris également des initiatives d'évaluation thématiques et de travaux méthodologiques. Enfin, les services de la Commission ont réalisé une série d'évaluations thématiques transnationales qui ont suscité un très grand intérêt.

À l'occasion de notre enquête, nous avons constaté un peu partout une appropriation gestionnaire de l'évaluation. Par exemple, en Finlande et en Irlande, les exigences européennes ont été fortement intégrées à la gestion des politiques nationales. En Suède, pays pourtant très accoutumé à l'évaluation, plusieurs autorités ont fait leurs premiers pas à l'occasion de la mise en œuvre des Fonds structurels, et se sont rapidement approprié la démarche dans une perspective managériale. Plusieurs régions d'Italie centrale ont également utilisé leurs évaluations à mi-parcours comme outil de gestion et de médiation.

L'évaluation comme acte politique

Au-delà de ce modèle managérial, l'évaluation peut aussi être conçue comme un outil du jeu démocratique, visant à informer le citoyen sur la raison d'être des dépenses publiques et sur leur efficacité, c'est-à-dire sur l'obtention des impacts

recherchés. Cette vocation était celle que préconisait le premier rapport soumis à Michel Rocard (Viveret, 1989).

Notre enquête a montré que la pratique européenne intègre rarement cette dimension. Seuls certains pays qui avaient une tradition préalable d'utilisation démocratique de l'évaluation, c'est-à-dire les pays scandinaves et le Royaume Uni, ont étendu cette conception, souvent à la marge, dans leur approche des Fonds structurels.

C'est au niveau européen que la demande d'utilisation « politique » de l'évaluation est la plus forte. Le Conseil et le Parlement insistent sans cesse pour que la Commission fasse preuve de plus de « redevabilité » (*accountability*) à propos de sa gestion d'une dépense qui dépasse maintenant 30 % du budget de l'Union.

La Commission publie donc périodiquement des rapports sur la mise en œuvre des Fonds et sur leur contribution à la cohésion économique et sociale. Ces rapports sont construits sur le principe que la politique européenne s'évalue en faisant la synthèse des évaluations régionales et nationales. En réalité, cet exercice de synthèse relève encore d'une gageure dans la mesure où la Commission doit réunir, dans une perspective politique, des matériaux dont la plupart ont été élaborés dans une perspective managériale.

Seconde conclusion : la culture d'évaluation reste fragile

Il est facile de comprendre qu'une utilisation démocratique de l'évaluation est particulièrement exigeante. Tant que l'évaluateur est considéré comme un conseil en management public ou comme un médiateur au sein du cercle des gestionnaires, les conséquences de l'évaluation restent « sous contrôle ». L'acceptation ou la non-acceptation des conclusions reste à la discrétion des gestionnaires. La pertinence des recommandations se discute en cercle fermé au sein de l'administration.

Cette pratique managériale conduit donc à privilégier les thèmes consensuels et les points de vue qui prédominent chez les gestionnaires. L'évaluation court alors le risque de se fondre dans les pratiques administratives traditionnelles.

Mes réflexions sur la construction des capacités d'évaluation m'ont conduit à penser que la pleine utilité de l'évaluation suppose, pour les gestionnaires, d'assumer pleinement le risque de conclusions négatives ou perturbantes (Toulemonde, 1999). Une telle attitude n'a aucune chance d'être spontanément adoptée à grande échelle. Elle ne peut être obtenue qu'à travers l'enracinement

d'une véritable culture de l'évaluation à tous les niveaux de l'administration. Elle nécessite également que les médias et les responsables politiques apprennent à se servir des travaux d'évaluation sans les dénaturer.

La création de sociétés ou d'associations nationales d'évaluation est un des moyens d'atteindre ce but. Les sociétés professionnelles d'évaluation éclosent dans de nombreux pays européens avec, par ordre approximatif de date de naissance, le Royaume-Uni, la Suisse, l'Allemagne, l'Italie, la France, et prochainement le Danemark, l'Espagne et la Finlande. Il existe aussi des sociétés ou clubs en Suède et depuis peu en Wallonie, ainsi bien sûr qu'une société européenne.

La Société d'évaluation britannique (UKES), est particulièrement active et semble pouvoir être considérée comme un modèle. Elle est forte de plus de 200 membres, organise une conférence annuelle et des ateliers de travail pour développer les échanges méthodologiques. Elle réunit des professionnels du secteur public et du secteur privé ainsi que des universitaires. Elle est un lieu de débats méthodologiques entre des professionnels dont l'expertise dans le domaine de l'évaluation est reconnue à l'échelle internationale. Le Royaume-Uni est un des rares États membres à disposer d'équipes de recherche en évaluation qui travaillent dans ce domaine de manière « transversale », sans se limiter à un aspect méthodologique particulier ni à un secteur politique donné.

La société britannique contribue aussi à faire admettre l'évaluation comme partie intégrante du fonctionnement démocratique. De fait, le Royaume-Uni est un des rares pays européens où l'administration utilise l'évaluation pour rendre des comptes aux citoyens.

Dans la majorité des États membres, la pratique de l'évaluation n'a pas du tout cette dimension, la culture politico-administrative n'y est pas favorable, et la pratique de l'évaluation reste donc incomplète.

6.4. Quelle place pour des modèles nationaux ?

Pour revenir sur la question posée au départ de cet article à propos de l'existence des spécificités nationales dans la pratique européenne, je conclurai que ces spécificités existent indéniablement aujourd'hui mais qu'elles ont vocation à disparaître.

Comme on l'a vu plus haut, de fortes différences sont repérables dans la façon dont l'évaluation s'est implantée et s'est développée d'un pays à l'autre. Des

différences encore plus grandes subsistent dans le degré d'intégration de l'évaluation dans la culture politico-administrative.

Par contre, il me semble qu'un mouvement de maturation est observable dans tous les pays européens. Cette maturation se caractérise par le fait que les initiatives d'évaluation sont de moins en moins concentrées dans les mains d'un champion national ou soumises à la pression extérieure.

Je définirais ce stade de maturité par le fait que l'évaluation s'applique à la plupart des politiques publiques, qu'elle est maîtrisée à tous les niveaux de gouvernement ou d'administration, qu'elle est pratiquée par des professionnels suffisamment polyvalents pour garantir un travail de qualité en toutes circonstances. Il me semble que beaucoup de pays européens ont atteint ce stade de maturité ou en sont proches.

Une autre forme de maturation est probablement à l'œuvre, qui voit les pays européens passer progressivement d'une évaluation subie à une utilisation managériale puis démocratique de l'exercice. Cette seconde forme de maturation est encore très loin d'être aboutie.

Mon espoir est que tous les pays de l'Union européenne parviendront au stade de maturité de leur pratique d'évaluation. Il me semble qu'à ce stade, quelques leçons universelles s'imposeront partout et prévaudront largement sur les spécificités nationales. Parmi les règles que je considère comme universelles figure l'unicité de chaque évaluation (Patton, 1986). Chaque situation d'évaluation mérite une réponse unique à des questions telles que : à quoi l'évaluation va servir ? quand doit-elle être faite ? qui doit y être impliqué ? quelles questions sont à poser ? quelle méthode doit être spécifiquement construite pour répondre aux questions ? On voit par là que la professionnalisation et l'acculturation à l'évaluation conduisent à des exercices « sur mesure » et non pas à des écoles ou à des modèles, qu'ils soient nationaux ou sectoriels.

Est-ce à dire qu'il n'y a pas de place pour des contributions spécifiquement nationales au corpus universel des connaissances en évaluation ? Je ne le pense certainement pas. En réalité, un travail d'évaluation de bonne qualité s'adapte très finement à la situation de la politique ou du programme soumis à examen, ainsi qu'au contexte politico-administratif de l'exercice. Cette adaptation pragmatique à la réalité devrait générer des évaluations multiformes et créer de très nombreuses occasions d'innover dans les façons de concevoir, de conduire ou d'utiliser les travaux. De fait, mes contacts dans les différents pays

européens m'ont fait découvrir une multitude de pratiques nouvelles, dont certaines sont tout à fait passionnantes.

Pour n'en citer qu'un exemple, je retiendrai l'utilisation de la technique des jurys de citoyens dans le cadre de nombreux travaux d'évaluation conduits au Danemark. Cette technique innovante est en train de s'exporter dans toute l'Europe et quelques tests ont déjà eu lieu en France. Il me paraît évident que cette innovation n'est pas née au Danemark par hasard. La culture démocratique y est en effet particulièrement ancienne et développée.

Sans parler de modèle danois de l'évaluation, ce qui serait à l'évidence exagéré, on peut cependant dire que les spécificités de la culture danoise ont créé une niche particulière, dans laquelle a germé une innovation prometteuse, et qui n'avait probablement pas de chances d'apparaître ailleurs. N'est-ce pas là une nouvelle preuve que la diversité des cultures nationales est une des principales richesses de l'Europe ?

6.5. Références bibliographiques

- Commission européenne, « Conception et conduite d'une évaluation », Collection MEANS, volume 1 Luxembourg , Office des Publications , 1999..
- Conseil scientifique de l'évaluation, « L'évaluation en développement 1998 »..La Documentation française , 1999.
- Furubo (J.E.), Rist (R), Sandhal (R), « International Evaluation Atlas », New Brunswick, Transaction Publishers, à paraître.
- Patton (M.Q), Utilization-Focused Evaluation, Newbury Park, Sage, 1986.
- Toulemonde (J.), « Incentives, Constraints and Culture-building as Instruments for the Development of Evaluation Demand », in Boyle (R.) et Lemaire (D.), « *Building Effective Evaluation Capacity* », New Brunswick, Transaction Publishers, 1999, p. 153-174.
- Viveret (P.). « L'évaluation des politiques et des actions publiques », rapport au Premier ministre, 1989..
- Wildawski (A.), « Sauver l'analyse de politique du PPBS », traduit de « Public Administration Review », « Politiques et Management public », 9 (4), 1969..

Contribution II

Les pratiques de l'évaluation française des contrats de Plan État-région

Par Maurice Baslé ¹

Introduction

L'expérience de la mise en œuvre partenariale de l'évaluation des contrats de Plan État-région pour la période 1993-1999 peut être analysée aujourd'hui et est propre à amener à une réflexion plus générale sur l'évolution des politiques partenariales en France et sur le développement de l'évaluation. On devrait en effet, par construction, théoriquement y détecter des terrains potentiels de coopération ou de médiation (Greffé, X., 1999) entre autorités officielles à différents niveaux de gouvernement et des occasions de pratiques nouvelles moins dépendantes d'un contexte propice à la compétition entre l'État et les collectivités territoriales. Les conditions de l'évaluation des contrats de Plan prévues par la circulaire du 9 décembre 1993 donnaient un cadre permettant l'expérimentation de cette coopération. La circulaire avait prévu que des crédits d'État à hauteur de 6/10 000^e des fonds engagés puissent être complétés par des crédits des conseils régionaux et financer le lancement d'études évaluatives coordonnées et suivies par des dispositifs régionaux d'évaluation (Baslé, M., 1998).

Avant d'analyser l'expérience et d'ouvrir la discussion (1 et 3), deux ensembles de rappels peuvent être faits. L'un sur le système politique français de superposition et d'emboîtement des décisions d'autorités officielles à divers niveaux (1) l'autre sur le développement des pratiques de l'évaluation et ses raisons d'être (2).

(1) *Maurice Baslé est professeur des universités.*

1 La superposition des aires de compétences et l'emboîtement des niveaux

Les politiques publiques sont décidées par des autorités officielles ayant une légitimité politique. Derrière la conception d'une politique publique, il y a évidemment une vision d'ensemble des « problèmes » à résoudre et une théorie de l'action, c'est-à-dire un ensemble de « modèles » plus ou moins cohérents et homogènes reliant directement ou indirectement des « causes » à des effets, les effets devant comprendre a minima quelques « solutions » aux problèmes identifiés au préalable. Ensuite, lors du « passage à l'action », le décideur place sa confiance dans l'action. Il a l'espérance qu'elle aura une influence (et la bonne influence). Ceci permet de dire qu'une politique publique est plus qu'une simple « annonce » : en outre, elle est une action singulière parce que c'est une action « politique » qui se situe très au-delà de l'action habituelle de la recherche de « l'intérêt », au sens d'un « intérêt particulier » (Buchanan, J., Congleton, R.D., 1998). Les politiques publiques, enfin, sont mises en œuvre par des « organisations en vue de l'action » obéissant à des pratiques de prise de décision, de suivi, de compte rendu et de bilan. On postule qu'il ne s'agit pas d'activisme (action pour l'action) mais bien d'une action orientée par des buts, objectifs officiels, sur la réalisation de laquelle le décideur doit rendre des comptes en terme de pertinence, cohérence, efficacité, efficience, effectivité et satisfaction des bénéficiaires. Sans oublier, au final, que derrière la politique publique, il y a la nécessité pour les élus, gouvernants à tous les niveaux de pouvoir et de gouvernement et leurs équipes, de faire comprendre ce qui est fait, de pouvoir communiquer de manière pédagogique sur l'action de préparer l'affichage d'un programme amélioré pour envisager sereinement la réélection et faire face aux sanctions du « marché politique ».

Ainsi sommairement résumées, les politiques publiques ont donc de multiples facettes (Muller, P., Surel, Y., 1999) : facettes symboliques (effets d'annonce, inaugurations, apparats...) ; facettes techniques avec les problèmes de gestionnaires (mise en place, suivi et contrôle) ; facettes techniques encore avec la question des choix multicritères ou encore avec les contextes d'incertitude qui pèsent nécessairement sur les politiques économiques et sociales (Cioffi-Revilla, C., 1998), facettes techniques aussi avec le respect des contraintes économiques et financières ; facettes politiques avec la question de l'évaluation et des interprétations socio-politiques (compréhension, communication, confrontations et gestion des conflits) et on n'oubliera évidemment pas la facette du contrôle.

Parmi les politiques publiques, au niveau le plus élevé de l'action politique, mais aussi à l'échelle des programmes et des sous-actions, chacun observe qu'on trouve aujourd'hui en Europe des actions qui ont quelque chose de singulier dans le paysage des politiques publiques : il s'agit de ces actions publiques qui sont complexes parce que contractualisées, polycentriques, polyarchiques, partenariales, associatives ou coopératives (par exemple, fonds structurels européens, contrats de Plan, contrats de pays, contrats de ville, schémas de services collectifs, etc.). Les raisons de ces nouvelles formes d'action et de complexité sont probablement d'abord la superposition des aires de compétences et la multiplication des véritables niveaux de pouvoir. Certes, le nombre d'autorités officielles, au sens strict, ne varie guère (les groupements de communes se développent cependant beaucoup) ; **certes, l'État reste en France, le poids lourd politiquement, administrativement et financièrement parlant.** Mais la multiplication des centres de décision et l'existence d'une réelle décentralisation en France nécessitent de plus en plus le besoin d'accords entre « niveaux de pouvoirs », donc une meilleure utilisation des compétences déconcentrées de l'État, et de meilleures négociations multi-niveaux, c'est-à-dire entre l'État, les collectivités territoriales et les institutions européennes. Ajoutons que par rapport à d'autres organisations et répartitions de compétences observables ailleurs ou théoriquement envisageables (Oates, W.E., 1972), le cadre français qui continue de prévaloir et qui est de plus en plus visible pour ce « gouvernement multi-niveaux », est un système de gouvernance particulier : on observe une spécialisation molle par bloc de compétences avec compétences exclusives et compétences partagées (Marchand, M.J., 2000 ; Guihery, L., 1997 ; du Granrut, C., 1997). Le principe de subsidiarité inscrit dans le traité de l'Union européenne permet en effet l'existence, en Europe, d'une grande variété historique et légitime de constitutions politiques, donc une grande richesse dans le fonctionnement multi-niveaux, hésitant entre le fédéralisme de type allemand (Guihery, L., 1997), la dévolution britannique, la décentralisation (Marchand, M.J., 2000) et la déconcentration française, le régionalisme catalan, etc.

Dans ce contexte français singulier, sur le fond, tout le monde peut comprendre qu'il y a une situation de jeux plus ou moins coopératifs. La donne peut être, pour simplifier, analysée en deux temps.

D'une part, les autorités officielles ont des histoires, des autorités, des légitimités inégales. De ce fait, même si elles souhaitent faire front commun et agir ensemble, elles sont inégalement dotées ce qui en fait des agents hétérogènes. Grosso modo, le budget de l'État représente en dépenses environ 1 800 milliards, celui des communes et groupements de communes (y compris la dette) environ 500 milliards, celui des départements environ 210 milliards,

celui de la contribution française aux politiques européennes est de 90 milliards environ et celui des régions françaises est de plus de 70 milliards. Dans le multi-niveaux, il y a des « grands pays » et « des petits » ou de manière plus réaliste, « il y en a qui sont moins égaux que d'autres ».

D'autre part, si les autorités officielles françaises de chaque niveau font souvent front commun avec quelques objectifs ultimes partagés (ce n'est pas nécessairement le cas de tous les objectifs de plus haut niveau, par exemple, l'Union européenne est, bien qu'émanation des 15 États membres, plus « européenne » que chacun des États membres), elles n'échappent pas à la concurrence entre autorités officielles sur certains objectifs ou pour certains champs (Breton, Y., 1996). Ce phénomène a longtemps été passé sous silence en raison de la prédominance, en France, de l'idée de la suprématie de la stratégie de l'État, de la centralisation financière et probablement, au plan politique, de la pratique française du cumul des mandats locaux et nationaux. De ce fait, certes, les différentes autorités officielles pratiquent la coopération par les financements croisés mais, en contradiction avec le principe coopératif qui les amène à co-financer et à « contracter », elles essaient aussi sur ces co-financements, de jouer « personnel » lorsque cela est possible et d'abord de mettre en avant leur contribution propre à l'action. Il arrive quelquefois aussi qu'elles aillent au-delà de leurs « périmètres » et de leurs prérogatives officielles écrites. Ainsi, les villes, districts, communautés, pays, départements, régions, les différents services déconcentrés de l'État et les différentes administrations centrales sont, en forçant le trait, autant « d'États en concurrence dans l'État » même si, en même temps, leurs liens, le tissu qu'ils constituent sont parfaitement réglementés et policés par les textes constitutionnels et dans des pratiques de jeux répétés. C'est une « famille nombreuse où l'on se chahute tout en restant en famille ».

La légitimité propre à chacun, l'inégal poids financier, la concurrence entre autorités officielles dont les pouvoirs se recouvrent, et le fait que les cibles des politiques publiques sont finalement les mêmes (territoires, bénéficiaires, électeurs) expliquent le souci de « partenariat quand même » pour conduire l'action et peuvent expliquer en même temps l'apparition de replis sur soi ou de méfiances malgré l'existence d'un terrain de médiation (Greffé, X., 1999).

2. Le développement de l'évaluation et les raisons de ce développement

Cette montée des politiques partenariales intervient en même temps qu'un autre souci se fait jour. De la même façon que l'action privée est de plus en plus évaluée sur les marchés, il est devenu de plus en plus important d'évaluer l'action publique et l'on observe un peu partout le développement des pratiques et celui des échanges d'informations grâce aux services rendus par les sociétés nationales d'évaluation ou des sociétés plus larges comme la Société européenne d'évaluation (Société française de l'évaluation, 1999).

Sans revenir sur l'histoire des méthodes de l'évaluation et sur les générations successives de débats doctrinaux (Baron, G., 1999), on peut rappeler la place spécifique de l'évaluation par rapport à la science, au contrôle, au conseil ou à l'audit. L'évaluation se situe du côté **du « comprendre », du « connaître », de l'analyse et de la préparation d'un jugement**. Rappelons qu'elle n'est pas assimilable à un contrôle, ni à un contrôle de régularité, ni à un contrôle administratif ni à une vérification comptable ni à un « contrôle de gestion », ni à une analyse financière. Rappelons qu'elle peut inclure l'audit des processus, des organisations ou des dispositifs impliqués dans la mise en œuvre de l'action évaluée ; qu'elle peut également utiliser des travaux d'études ou de recherche de caractère purement analytique. Rappelons également que l'évaluation ne peut cependant se réduire à ces derniers procédés éléments de connaissance, car il s'agit de s'approcher du point de vue de la commande de la politique et de mettre en œuvre ce qui va permettre de comprendre la politique étudiée dans son ensemble, d'avoir une vue complète de ses effets, d'apprécier globalement le degré d'atteinte de ses objectifs et l'efficacité dans l'usage des ressources publiques qui lui sont affectées. (Conseil scientifique de l'évaluation, 1996). L'évaluation met en œuvre un processus de construction de savoir même si l'évaluation ne peut se réduire à aucune modalité ou technique spécialisée de connaissance ou d'analyse, car elle consiste précisément en la mise en œuvre d'une médiation qui permet le croisement et la mise en débat des points de vue et des approches cognitives d'un objet spécifique : les politiques publiques et leur conduite. À cet égard, il est spécifique à l'évaluation d'être en même temps connaissance ouverte car il n'y a pas d'évaluation sans expertise indépendante et regard extérieur sur la politique évaluée, mais il est aussi spécifique de l'évaluation d'encourager la prise en compte privilégiée des questions, besoins et critères de jugement des acteurs, décideurs et bénéficiaires (Abma, T.A.,

1997). Ce pluralisme, qui se place idéalement à égale distance de l'auto-évaluation, de l'expertise maîtrisée par le commanditaire et du jugement scientifique pur par une autorité indépendante, va prendre la forme **d'une instance d'évaluation indépendante**, garante de l'expression des divers points de vue légitimes en même temps que de la qualité et de l'honnêteté du travail scientifique servant de base au jugement évaluatif.

Trois ensembles de causes ont probablement contribué au développement du réflexe et de la culture de l'évaluation des politiques publiques dans les pays les plus développés et dans les démocraties les plus ouvertes depuis une quinzaine d'années et en particulier dans les quinze pays de l'Union européenne et au niveau de l'Union européenne et, en particulier, pour des politiques ayant des effets sur des territoires infra-nationaux.

Première raison, les contraintes financières pèsent de plus en plus sur les budgets publics aux différents niveaux de gouvernement. Ces contraintes ne sont pas seulement la conséquence conjoncturelle du ralentissement de la croissance en Europe au cours des années 1990 : elles présentent désormais un caractère structurel, lié à la mondialisation de la concurrence. La baisse des prélèvements obligatoires (même nets des transferts aux entreprises et du financement d'infrastructures utiles pour la compétitivité), devient une variable d'action pour les gouvernements placés sous la contrainte de leurs opinions publiques, elles-mêmes soumises au moins partiellement aux pressions d'une culture de la concurrence mondiale. La concurrence fiscale des territoires participe de ces contraintes.

Deuxième raison, une raison politique. On demande plus aux marchés, on demande plus aux États et aux collectivités territoriales. Cette pression exercée par l'opinion trouve donc ses racines dans **l'évolution et l'enracinement de la culture démocratique**. Tout d'abord, et ça n'est pas nouveau, le citoyen devient de plus en plus exigeant et attend de la collectivité qu'elle règle les difficultés qu'il rencontre. Tout en regrettant le nombre, excessif selon lui, de fonctionnaires, il demande toujours plus de service public et au service public. Avec l'accroissement des niveaux d'éducation, les citoyens et les opinions publiques souhaitent comprendre davantage ce qui est fait de l'argent public. Il souhaite mieux connaître les effets de l'action publique, être à même de vérifier si l'on tient ce qu'on promet (efficacité dans la réalisation des objectifs officiels), si on limite le gaspillage (efficacité), si on utilise les bons outils (cohérence), si les politiques sont adaptées aux problèmes à résoudre (pertinence). Il souhaite être mieux informé et mieux peser là où il a quelque poids, à savoir sur les « marchés politiques » aux différents niveaux de

gouvernement (du communal au niveau européen). Les risques de censure ou de non-réélection, signes de bonne démocratie, influencent de manière croissante la conduite des élus et des administrations en place. En réponse, ceux-ci doivent connaître (évaluation), faire connaître (fournir des connaissances appropriables), comprendre et faire comprendre (être crédibles) car on ne peut guère communiquer sans substance (Patton, M.Q., 1997). Le principe de subsidiarité en même temps qu'une certaine confiance dans des mécanismes décentralisés ou de proximité renforcent la pression politique sur les instances locales.

Troisième raison, la gestion publique. Le développement de l'évaluation s'explique aussi comme élément de réponse à de nouveaux **défis managériaux en interne, défis qui ne sont pas propres au secteur public**. En effet, si l'entreprise privée développe de plus en plus une stratégie de management dont l'objectif central est la satisfaction du client potentiel, véritable censeur de la qualité du produit dans une économie de marché, il appartient aussi de plus en plus aux services publics de replacer au cœur de leur système d'action les usagers bénéficiaires de leurs prestations. Le retard en la matière peut s'expliquer par diverses causes. Une de ces causes est que l'utilisateur du service public est un client captif, qui n'a pas la liberté du choix et a rarement celle de donner son appréciation sur la qualité du service dont il a bénéficié. Le changement, c'est qu'avec l'éducation croissante des citoyens, le « client » change. C'est pour cette raison que les services administratifs aux différents niveaux de gouvernement ont maintenant besoin de faire évoluer leurs systèmes d'information et de gestion physique et financière, de mieux piloter leur action, de créer les outils adaptés pour ce faire (tableaux de bord). Les besoins d'ingénierie de l'action publique se font jour. Ceci a des conséquences sur les besoins en formation continue de personnels de différentes catégories car on ne peut faire évoluer un processus, la mise en œuvre d'un programme ou d'une politique économique sans développer les connaissances et les bons réflexes des gestionnaires dans les services (administrations centrales, services déconcentrés de l'État, agences publiques, établissements publics, collectivités territoriales...). On ne peut infléchir un programme sans avoir d'idées des effets de l'action engagée et sans disposer de programmes de rechange. Les politiques publiques ont donc aussi besoin de management stratégique, c'est-à-dire de plusieurs procédures : d'une information large et diversifiée, d'un processus de sélection de cette information et de dispositifs permettant d'élaborer des jugements de valeur sur l'action menée. Ceci vaut tout particulièrement pour la fonction publique territoriale et les agents des services déconcentrés de l'État qui sont en rapport direct avec des « usagers » locaux.

3. Le cas de la mise en œuvre partenariale des politiques publiques : les contrats de Plan État-région en France de 1993 à 1999

L'expérience des contrats de Plan État-région 1993-1999 et de leur évaluation s'achève. Les politiques de ces contrats ont été évaluées dans les conditions prévues par la circulaire du 9 décembre 1993. Le financement prévu par la circulaire faisait mention de crédits d'État affectés à hauteur de 6/10 000 des fonds engagés ce qui a donné des moyens aux SGAR (secrétaires généraux d'administration régionale) pour entreprendre seuls ou en partenariat, des évaluations de programmes.

L'observation montre que de nombreux chantiers d'évaluation ont été ouverts sur la période 1993-1999 avec une préférence de fait pour des politiques que les autorités officielles des différents niveaux de gouvernement portent souvent en commun, en l'occurrence dans le cas qui nous occupe, les politiques de la ville, les politiques en faveur de l'agriculture et/ou de l'environnement, les politiques d'aides à l'emploi et au développement économique, les aides aux entreprises et aussi les aides aux transferts de technologies.

3.1. Les pratiques partenariales de l'évaluation des contrats de Plan État-région (1993-1999)

On ne donne ici que deux exemples et des éléments synthétiques.

a) Politiques de la ville

Des sources existantes d'informations ont été utilisées et des co-productions d'indicateurs ont eu lieu qui témoignent de la volonté d'évaluation partenariale (tableau 1 et tableau 2).

b) Politiques d'aide aux entreprises

La même remarque s'impose pour l'utilisation commune de nouvelles sources d'informations (tableau 3) et le choix des impacts étudiés en commun (tableau 4).

Tableau 1 : Source des informations généralement retenues pour l'évaluation de la politique de la ville. (synthèse)

Évaluation de l'emploi et de l'insertion par l'économique.	Évaluation de l'accessibilité des services publics dans les quartiers prioritaires.	Évaluation de l'habitat, transports et urbanisme.	Évaluation de l'éducation et de l'environnement scolaire.	Évaluation de la sécurité et de la prévention de la délinquance.
<ul style="list-style-type: none"> • <u>Insertion par l'économique</u> : entretiens auprès des acteurs institutionnels. • <u>Chômage</u> : données du fichier individuel de l'ANPE, exploitées par l'INSEE. 	<ul style="list-style-type: none"> • <u>Action sociale</u> : entretiens auprès des travailleurs sociaux. 	<ul style="list-style-type: none"> • <u>Habitat et aménagement urbain</u> : entretiens auprès des bailleurs sociaux et acteurs institutionnels ainsi que des enquêtes auprès des usagers et données recueillies auprès des organismes HLM et de l'Association Régionale HLM. 	<ul style="list-style-type: none"> • <u>Éducation</u> : entretiens auprès de l'Education nationale, de la DDASS et de la CAF. 	<ul style="list-style-type: none"> • <u>Prévention-sécurité</u> : données statistiques fournies par les services concernés : données de la Direction régionale de la protection judiciaire de la jeunesse et de la DDPJJ et dossier ressources de la DIV.
<ul style="list-style-type: none"> • <u>Précarité</u> : données de la CAF et de l'INSEE. • <u>Emploi</u> : données de l'ANPE-ASSEDIC, de la DRTEFP et de la DRFAS, dossier ressources de la DIV et entretiens auprès des acteurs de l'emploi et de l'insertion, et principalement de l'ANPE. 	<ul style="list-style-type: none"> • <u>Santé</u> : réunions avec les services compétents, entretiens et données de la DRASS et dossier ressources de la DIV. • <u>La Poste</u> : entretiens. 	<ul style="list-style-type: none"> • <u>Désenclavement et aménagements urbains</u> : entretiens auprès des habitants. 	<ul style="list-style-type: none"> • <u>Ecole</u> : informations fournies par le Rectorat et entretiens auprès de plusieurs catégories de personnel de l'Education Nationale (IA, chargés de mission, responsables de service) et données chiffrées sur l'utilisation des fonds versés au titre de la politique de la ville. 	<ul style="list-style-type: none"> • <u>Police</u> : entretiens auprès de différents acteurs, agents de la police nationale et auprès des chefs de projet des sites en contrat de ville.
<ul style="list-style-type: none"> • <u>Aide à la requalification des centres commerciaux</u> : coût et financement des opérations de restructuration pour les différents sites choisis. 		<ul style="list-style-type: none"> • <u>Habitat</u> : données fournies par les fichiers sur le parc locatif social, gérés par la DRE et l'AROHLM et dossier ressources de la DIV. 	<ul style="list-style-type: none"> • <u>Ecole et son environnement</u> : données de l'IA et des Directeurs d'établissements, et données de la DRAC et de la DRJS. 	

Tableau 2 : nature des indicateurs généralement retenus pour l'évaluation des politiques de la ville.

Évaluation de l'emploi et insertion par l'économique.	Évaluation de l'accessibilité des services publics dans les quartiers prioritaires.	Évaluation de l'habitat, transports et urbanisme.	Évaluation de l'éducation et de l'environnement scolaire.	Évaluation de la sécurité et de la prévention de la délinquance.
<ul style="list-style-type: none"> • <u>Précarité</u> : indicateurs quantitatifs. • <u>Chômage</u> : indicateurs quantitatifs. 	Pas d'utilisation d'indicateurs.	<ul style="list-style-type: none"> • <u>Logement, parc locatif social</u> : indicateurs quantitatifs. 	<ul style="list-style-type: none"> • <u>Education</u> : indicateurs quantitatifs. 	<ul style="list-style-type: none"> • <u>Prévention de la délinquance-sécurité</u> : indicateurs quantitatifs.
<ul style="list-style-type: none"> • <u>Emploi</u> : indicateurs quantitatifs. • <u>Activité</u> : indicateurs quantitatifs. <ul style="list-style-type: none"> • <u>Insertion par l'économique</u> : indicateurs quantitatifs. 		<ul style="list-style-type: none"> • <u>Aide à la requalification des centres commerciaux</u> : indicateurs quantitatifs et qualitatifs. • <u>Habitat, urbanisme, transports</u> : indicateurs quantitatifs et qualitatifs. 	<ul style="list-style-type: none"> • <u>Education</u> : indicateurs du premier type plutôt quantitatifs et indicateurs du second type , plutôt qualitatifs. 	

Tableau 3

Sources d'informations qualitatives et quantitatives généralement retenues pour l'évaluation des politiques en faveur des entreprises.

Sources d'informations retenues.	
Qualitatives	Quantitatives
<ul style="list-style-type: none"> ➤ Enquête approfondie en face-à-face auprès d'un échantillon de chefs d'entreprises. ➤ Entretiens approfondis auprès des services instructeurs de l'Etat et de la région. 	<ul style="list-style-type: none"> ➤ Compte d'exploitation et bilan des entreprises bénéficiaires.
<ul style="list-style-type: none"> ➤ Enquête auprès des bénéficiaires. ➤ Enquête auprès des services instructeurs. 	<ul style="list-style-type: none"> ➤ Données statistiques issues de la base de données de la DRIRE. ➤ Fichiers obtenus auprès du SESSI. ➤ Dossiers d'aides. ➤ Liasses fiscales complètes.
<ul style="list-style-type: none"> ➤ Enquête auprès des bénéficiaires. ➤ Enquête auprès des non-bénéficiaires. ➤ Interviews auprès de bénéficiaires et de non-bénéficiaires, de gestionnaires des aides, d'une société de conseil et de quelques autres personnalités. 	<ul style="list-style-type: none"> ➤ Fichier GESTCRAE. ➤ Publications économiques ➤ Base de données SIRENE.
<ul style="list-style-type: none"> ➤ Enquête auprès des consommateurs. ➤ Enquête auprès des bénéficiaires. ➤ Enquête auprès des non-bénéficiaires. ➤ Guides d'entretiens (aux élus du site et à l'association des artisans et commerçants). 	<ul style="list-style-type: none"> ➤ Dossiers de demandes de subventions.

<ul style="list-style-type: none">➤ Enquête auprès des chefs d'entreprises de la région (par la DDER).➤ Enquête de l'INSEE/ Direction de l'artisanat auprès des PEI pour les entreprises de petite taille.➤ Entretiens auprès des différents acteurs extérieurs aux services de l'Etat et de la région directement impliqués dans la procédure d'aides.➤ Comptes rendus des réunions du groupe mixte industrie (service des études du SGAR).	<ul style="list-style-type: none">➤ Statistiques du fichier SIRENE.➤ Informations sur les bilans des entreprises (Eurydile, sur Minitel).➤ Informations contenues dans les protocoles des entreprises qui ont contracté avec la région dans le cadre du dispositif.
---	---

Tableau 4

Impacts qui ont été étudiés lors de l'évaluation des politiques d'aides en faveur des entreprises et leur nature

Nature des impacts.	Impacts étudiés.
Impacts déclarés par les entreprises bénéficiaires et impacts observés, à l'aide du compte d'exploitation et du bilan.	Impact sur le plan de financement des entreprises.
Impacts déclarés par les entreprises bénéficiaires et impacts observés, à l'aide des éléments financiers obtenus à partir des liasses fiscales et informations complémentaires récoltées auprès des entreprises lors de l'enquête.	Impacts sur l'évolution générale de l'entreprise, sur l'évolution de la compétitivité, de la réactivité et la flexibilité ainsi que sur l'évolution de la compétence du personnel.
Impacts déclarés par les entreprises bénéficiaires.	Impact sur la nature de l'activité des entreprises, sur l'économie des entreprises et sur l'emploi.

Impacts déclarés par les entreprises bénéficiaires et les consommateurs. Impacts observés, à l'aide des factures, et impacts déclarés par les consommateurs. Impacts déclarés par les bénéficiaires. Impacts déclarés par les bénéficiaires.	Effets micro (sur les bénéficiaires, les comportements des consommateurs et sur l'investissement (I) immatériel). Effets macro (retombées économiques locales de l'I : emplois, et retombées en termes de qualité de vie). Mise en œuvre du dispositif. Appréciation de la procédure.
Impacts observés, à partir des données relatives à chaque type d'aides. Impacts posés comme hypothèse ¹ .	Effets individuels. Incidence économique sur les activités et les marchés. Impact macro au niveau régional sur la production, l'emploi, le développement.

3.2. Les organisations nécessaires à l'évaluation : instances et dispositifs partenariaux

Diverses instances plus ou moins permanentes, plus ou moins partenariales, plus ou moins externes ont été créées, ont fonctionné quelquefois avec difficulté, dans les différentes régions (Baslé, M., 2000, chapitre I). Les informations disponibles permettent de dresser une première liste théorique de tous les dispositifs régionaux en 1998 (tableau 5).

¹ Les évaluateurs ont mené des réflexions afin de voir quels sont les impacts qui pourraient être observés en théorie et proposent ainsi des sources d'informations sur lesquelles il faudrait s'appuyer pour les analyser.

**Tableau 5 Récapitulatif et bilan des institutions
et modes de “portage” des instances (1998) ***

Régions	Organisation
Alsace	Groupe de travail Etat-région
Aquitaine	Comité d'évaluation et bureau du comité de pilotage, groupes de suivi ad hoc
Auvergne	Comité Etat-région avec instances techniques ad hoc pour le suivi
Basse-Normandie	Comité Etat-région avec instances techniques et groupes de travail
Bourgogne	Comité Etat-région avec comités techniques ad hoc
Bretagne	Comité consultatif Etat-région-départements-conseils généraux, Comité régional Etat-région, Commission scientifique et institutionnelle, secrétariat permanent, Groupes de suivi. Participation à l'évaluation des contrats de ville et des Fonds structurels. Association
Centre	Comité de pilotage Etat-région. Des groupes ad hoc de suivi
Champagne-Ardenne	Comité régional Etat-région. Comité de pilotage. Groupes techniques sectoriels.
Corse	Comité de pilotage. Instances techniques ad hoc. Experts.
La Réunion	Comité de pilotage. Instances techniques.
Franche-Comté	Comité régional Etat-région. Collège scientifique. Groupes de suivi ad hoc.
Haute-Normandie	Comité régional de suivi et d'évaluation Etat-Région. Comités techniques.
Ile de France	Comité régional d'évaluation Etat-région (et de suivi). Instances techniques à géométrie variable. Rapporteur comme secrétaire de l'instance.
Languedoc-Roussillon	Comité régional d'évaluation avec un Président, le président du Conseil économique et social régional (CESR), 9 membres du CESR, 10 experts. Groupes techniques ad hoc (Etat).
Limousin	Comité de pilotage de l'évaluation Etat-région. Comité technique SGAR-Services de la Région-personnalités qualifiées. Groupes de travail avec experts.
Lorraine	Comité de pilotage de l'évaluation Etat-région-TPG-université. Comités techniques avec experts. Cellule associative de portage pour l'animation.
Midi-Pyrénées	Comité régional d'évaluation Etat-région-CESR-Banque de France-TPG. Commission scientifique présidée par le président de la Chambre régionale des comptes. Groupes de projet et de suivi des chargés d'évaluation.
Nord-Pas-de-Calais	Comité de suivi du Contrat de plan. Comité de pilotage de l'évaluation avec experts. Cellule d'évaluation Etat-Région. Groupes thématiques.
Provence-Alpes-Côte-d'Azur	Comité de pilotage de l'évaluation Etat-région. Délégation des responsabilités au Président de la commission des Finances du plan et de l'évaluation du Conseil régional.
Pays de la Loire	Comité régional d'évaluation Etat-région. Commission scientifique. Groupes de projet et de suivi.
Picardie	Comité régional d'évaluation Etat-région-CESR-CRC-INSEE-BdF-TPG-Université. Instances techniques pluridisciplinaires.
Poitou-Charentes	Comité régional d'évaluation Etat-région-CESR-TPG-CRC-un expert. Groupe techniques. Secrétariat permanent de l'Association Institut Atlantique d'Aménagement du Territoire.
Rhône-Alpes	Comité régional de suivi-évaluation Etat-région-CESR-personnalités qualifiées. Comité régional scientifique avec secrétariat permanent. Groupes de suivi.

- Documentation mise à jour en 1998 par Karine Lagant pour le Comité régional d'évaluation de Bretagne.
-

Les dispositifs peuvent d'abord être classés de façon à faire apparaître de grandes différences ou de grandes similitudes. Sur ce point, les principaux organes qui sont à l'origine de dispositifs d'évaluation originaux sont au nombre de deux. On citera tout d'abord le comité scientifique et ensuite la présence ou non d'un secrétariat permanent comme il peut en exister en région Bretagne ou en région Poitou-Charentes. Dans d'autres cas, comme en Aquitaine (avec le bureau du comité de pilotage) ou en Limousin (avec un groupe de travail), les distinctions apparaissent sous des formes plus spécifiques, issues de la volonté d'organiser un dispositif selon un autre schéma.

En matière d'évaluation, certaines régions sont aux antipodes. On citera par exemple le cas de l'Alsace qui ne disposait pas en 1998 de dispositif régional permettant de mener à bien les évaluations (mais qui travaille dans ce sens), et la région Nord-Pas-de-Calais.

L'importance des travaux effectués en évaluation n'est cependant pas en liaison directe avec la taille du dispositif mis en place. Dans certaines régions, les avancées sont considérables sans pour autant que les organes intervenant dans ce travail soient nombreux, le but étant en effet de ne pas multiplier les procédures.

On peut aussi regarder l'existence d'un comité scientifique. Son apport est très important mais très peu de régions ont mis en place une telle instance, préférant intégrer ou convoquer ponctuellement des experts dans des groupes ou des comités préexistants. Le monde d'utilisation des experts et leur statut dans le processus de l'évaluation est évidemment très inégal et ceci est préoccupant à bien des regards : indépendance, qualité, professionnalité doivent être le plus possible préservées et les fonctionnements des dispositifs ne peuvent pas toujours le garantir. Sur tous ces points que l'on pourrait détailler pour certaines régions (Baslé, M., 1998 ; Isaïa, H., 1999), de nouveaux travaux pourraient être commandés.

Aujourd'hui, formuler des conclusions plus élaborées semble inutile et difficile. En effet, l'évaluation est d'abord une activité nouvelle (pour tout ce qui concerne les évaluations du contrat de Plan État-région) rendue obligatoire depuis la circulaire de décembre 1993. L'évaluation des politiques publiques (certes en plein essor) s'opère à un rythme plus ou moins rapide selon les régions, et aussi selon la volonté des acteurs. On peut observer aussi que le non-respect de la circulaire de 1993

n'entraîne aucune sanction. L'évaluation avait par ailleurs débuté dans certaines régions avant même que soit mise en œuvre la circulaire, ce qui a permis à ces régions de conserver une certaine avance et de disposer d'une plus grande expérience.

Un élément comparatif au dispositif proposé par les circulaires peut être apporté. On peut comparer certains dispositifs à une situation « légale » ou « référentielle » décrite dans les grands textes de loi relatifs à l'évaluation ; notamment par la circulaire du 9 décembre 1993. La comparaison débouche simplement sur trois formes qui conduisent à la typologie suivante :

- les régions qui n'ont pas suivi les propositions institutionnelles minimales,
- les régions qui se sont cantonnées aux propositions institutionnelles,
- les régions qui sont allées plus loin que les propositions institutionnelles.

Les régions ont ainsi été classées en trois catégories dans le tableau de synthèse suivant (tableau 6)

Tableau 6

Dispositifs observés et dispositif de référence pour l'évaluation des CPER (1993-1999)

Régions plus faiblement dotées que ne le prévoit le dispositif institutionnel	Régions dont le dispositif est semblable au dispositif institutionnel	Régions dont le dispositif est plus important que le dispositif institutionnel
Alsace Provence-Alpes-Côte d'Azur	Auvergne Basse-Normandie	Aquitaine Bretagne
	Bourgogne Corse DOM TOM Haute-Normandie Ile-de-France Languedoc-Roussillon Picardie	Centre Champagne-Ardenne Franche-Comté Limousin Lorraine Midi-Pyrénées Nord-Pas-de-Calais Pays de la Loire Poitou-Charentes Rhône-Alpes

On peut aussi regrouper les régions en sous-ensembles sur la base du management de l'évaluation on compare les compositions du comité de pilotage de l'évaluation (appelé aussi comité régional). On a pris en compte, pour toutes les régions, la totalité des membres qui figurent dans le dispositif d'évaluation (peu important qu'ils fassent partie ou non du comité de pilotage). Les régions peuvent alors être réparties en trois grandes catégories (tableau 7)

Tableau 7 - Comparaison des comités de pilotage pour l'évaluation des CPER

Régions plus faiblement dotées en membres que ne le prévoit le dispositif légal	Régions dont le nombre de membres est semblable au dispositif légal	Régions dont le nombre de membres est plus important que le dispositif légal
Bourgogne	Aquitaine	Bretagne
Centre	Auvergne	Franche-Comté
Haute-Normandie	Basse-Normandie	Limousin
Languedoc-Roussillon	Champagne-Ardenne	Midi-Pyrénées
	Corse	Nord-Pas-de-Calais
	Ile-de-France	Pays de la Loire
	Lorraine	Rhône-Alpes
	Picardie	
	Poitou-Charentes	

On observe que de nombreuses régions ont respecté les dispositions légales ou ont tenté de faire mieux. Toutefois, pour certaines, des précisions doivent être fournies : certaines d'entre elles ont été considérées comme pourvues de membres de façon équivalente à ce que préconisait la circulaire du 9 décembre 1993, alors que si l'on établit une stricte comparaison, on remarque que certaines personnes n'interviennent pas dans le dispositif.

Au total, en matière de comparaison avec le dispositif institutionnel de référence, on observe des choix régionaux plutôt diversifiés même si un grand nombre de régions ont au minimum respecté le dispositif exposé dans la circulaire de 1993. Ce trait est véritablement encourageant au vu des nombreuses difficultés d'installation d'un tel dispositif. Alors qu'aucune sanction n'existe à l'encontre des régions qui ne respectent pas les propositions

énoncées par les circulaires, il est intéressant de remarquer que la plupart des régions ont agi conformément à la circulaire de 1993. Les trois quarts d'entre elles ont fait encore mieux. Reste à considérer si le dispositif a ouvert ou non la voie à des procédures réellement coopératives entre l'État et les régions

4. Réalisation coopérative des évaluations de contrats de Plan État-région (1993-1999)

Un bilan rapide des travaux menés montre des succès incontestables en même temps que des réalisations des évaluations pas toujours parfaites. L'ambition a été, le plus souvent, d'utiliser les meilleurs outils, de réunir les meilleurs spécialistes et de les rapprocher des services et d'assurer les conditions de l'indépendance de la production d'une information fiable, utilisable et partagée (Isaia, H., 1999, p. 71). Les expériences n'ont réalisé cette ambition qu'en tendance avec un ensemble de bonnes pratiques tant en processus qu'en mobilisation d'experts indépendants et de qualité. Plus inégale a été, entre l'État et les régions, la mise en commun de l'évaluation : on observe dans quelques cas seulement les pratiques régulières et communes de création de cahiers des charges, d'appels d'offres, d'animation de comités de suivi, d'audition des experts, de discussion des résultats devant les services, et celles-ci provoquent alors dans de nombreux cas des avancées significatives, du moins quand les relations entre les services déconcentrés de l'État en région et les services correspondants du conseil régional ont accepté de se pré-coordonner, de faire preuve d'un minimum de volonté d'aboutir et ont créé les conditions du développement de la confiance mutuelle. Il a fallu qu'il y ait à la fois la démarche constructive d'aller vers l'autre, de le reconnaître pour ce qu'il est, de savoir que l'autre cherche à savoir s'il est reconnu pour ce qu'il est et le partage d'un minimum de culture commune dans l'évaluation et de foi commune dans l'utilité de ses résultats. Un certain professionnalisme, beaucoup de réalisme, beaucoup de pragmatisme, du courage politique, ont facilité dans certains cas le minimum d'accord préalable sur les mots, les méthodes, les finalités et les procédures de résolution des désaccords. (Comité régional d'évaluation en Bretagne, 1999). Dans le cas de la Bretagne, il a été nécessaire que les jeux répétés de l'évaluation partenariale en région donnent lieu à des anticipations de surplus collectif avec un intérêt pour chacun. Il a été nécessaire que chacun ait pu aussi envisager la capitalisation de la connaissance produite mais aussi la capitalisation de la culture de l'évaluation réalisant ainsi en même temps des apprentissages individuels et un apprentissage organisationnel (sur ce point, voir l'exemple des fonds européens dans Baron, G., 1999).

Dans le cas de la Bretagne les outils, les méthodes ont été partiellement transférés d'un domaine à l'autre. Les informations quantitatives et qualitatives ont pu alors s'enrichir mutuellement. Ceci a été particulièrement le cas dans l'expérience d'évaluation dite légère (mais exhaustive) réalisée en Bretagne (expérience qui n'est certes pas parfaite). Les apprentissages de procédures de construction des indicateurs s'y sont avérées transférables d'un domaine à l'autre. Plus généralement, l'informatisation des tableaux de bord se fait une fois pour toutes (pour un résumé des résolutions de problèmes de tableaux de bord dits d'évaluation légère du contrat de Plan État-région Bretagne, voir Comité régional d'évaluation, 1999).

Conclusion

Au bout du compte, la leçon des évaluations partenariales des contrats de Plan État-région 1993-1999 est un constat d'humilité. Quelque chose a bougé dans nombre de relations entre les SGAR et les Conseils régionaux. L'évaluation est un champ d'études et de médiation, un espace ou un terrain de jeux stratégiques et en même temps, une possibilité d'apprentissage et de valeur ajoutée pour préparer des jugements politiques, des inflexions de programmes et des partages des débats.

Du point de vue de l'utilisation de l'évaluation, des conséquences en termes de formation interne, de management public partagé sont également observables. Surtout, on a l'impression que ce qui a été fait, pour quelques services et quelques programmes au moins a permis de préparer les services aux négociations de la nouvelle génération des contrats de Plan (2000-2006). Les votes concernant cette nouvelle génération sont à l'évidence des votes politiques au plus haut degré. Mais il est probable que l'instruction et la négociation nouvelles ont bénéficié, dans certains cas, de l'aiguillon du principe de l'évaluation et des résultats des évaluations précédentes. En Bretagne, un texte interne a été produit qui est une note « pour rendre plus et mieux évaluable » le nouveau CPER. L'écriture des conventions doit mieux soigner la présentation des objectifs et envisager les « zones à évaluer » et la production de systèmes d'indicateurs. La nécessité d'une approche partenariale de l'évaluation de politiques à financements croisés apparaît encore plus au grand jour. Cette apparition a certes été très lente. Mon sentiment personnel est que le processus est cependant inéluctable. Pour diverses raisons.

J'insisterai, en conclusion, surtout sur la crédibilité des commanditaires en même temps que leur besoin d'assurance mutuelle contre les risques d'échec des programmes ou des politiques. On comprend en effet facilement que la crédibilité des travaux d'évaluation sur des politiques en région est tout simplement plus grande dès lors qu'ils sont menés en commun (Robson, C., 2000, chapitre 2). Un préfet de région qui doit, de toute façon, procéder à des évaluations au titre de la circulaire du 9 décembre 1993, a quand même tout intérêt à créer son club régional et à faire participer les directeurs de service du conseil régional et les experts : pour certaines actions, tout d'abord, c'est le moyen d'obtenir, de vérifier ou de compléter son information ; pour d'autres actions ensuite, c'est le gage qu'il ne s'agit pas d'un suivi purement interne ou d'une auto-évaluation (c'est-à-dire de quelque chose qui n'est pas une évaluation). Plus généralement, c'est une ouverture à des « autorités officielles autonomes » mais évoluant sur les mêmes territoires. Inversement un président de conseil régional, aussi autonome soit-il, n'a pas souvent les moyens de procéder unilatéralement. En outre, la même remarque peut lui être faite concernant la nécessité d'éviter l'auto-évaluation. La mise en œuvre volontariste des conditions de respect, confiance et réciprocité permet en cas d'évaluation partenariale, d'envisager sereinement que le jeu restera bien un jeu coopératif. On peut difficilement envisager que la connaissance des résultats d'une action commune puisse se faire par appel séparé à des experts ou aux gestionnaires, chacun essayant de se former son jugement de valeur propre sur l'effectivité ou la pertinence d'une politique ou d'un programme. Au contraire, une meilleure mise en commun de la technologie de l'évaluation, une culture partagée sinon commune dans l'entreprise « évaluation » offre des garanties supplémentaires pour la connaissance et donc la performance publique. C'est, en outre, **un pacte évaluatif** qui est une véritable assurance mutuelle. On le voit, malgré la compétition forte entre les autorités officielles à divers niveaux de gouvernement, dès lors que le principe de subsidiarité et la tradition française créent une organisation complexe et encore hiérarchique des politiques, des zones de coopération ont été trouvées malgré les apparences dans une solution qui paraît a priori compliquée et déroutante : les financements croisés de politiques et de programmes et leur évaluation partenariale pour la période 1993-1999. La période nouvelle qui s'ouvre (2000-2006) avec la coopération renforcée dans l'Union européenne, avec la préparation de l'élargissement, avec aussi probablement une évolution (dont on dessine mal les contours mais qui est absolument nécessaire) de la déconcentration des services de l'État (renforcement du caractère interministériel des préfets ?) et de la décentralisation (demandes de statuts régionaux à la carte ?) sera probablement l'occasion de voir se développer les tendances (en partie chaotiques, en partie maîtrisées) au développement des dispositifs et des « bonnes pratiques » de

l'évaluation des politiques publiques territoriales. En espérant qu'on ne lira pas cette affirmation comme un pur acte de foi mais comme le résultant d'un vrai bilan des expériences humaines et sociales riches qui ont été menées à bien jusqu'à maintenant !

Bibliographie

- Abma, (T.A.), « Playing with and in Plurality », *Evaluation*, volume 3, n° 1, 1997, p. 25-48.
- Acocella, (N.), *The Foundations of Economic Policy Values and Techniques*, Cambridge University Press, 1998.
- Aristigueta, (M.P.), *Managing for Results in State Government*, Quorum Books, 1999.
- Baron, (G.), *Évaluation, participation, apprentissage : une conception de l'action publique avec rationalité limitée*, thèse de doctorat en sciences économiques de l'université de Rennes-1, décembre, à paraître, 1999.
- Baron, (G.), « Impact of Stakeholders Participation in a Mid-term Evaluation on Individual and Organizational Learning : a Case Study », *Evaluation*, Sage, 2000 à paraître.
- Baslé, (M.), *Le budget de l'État*, collections Repères, La Découverte, 1997, (nouvelle édition en préparation, à paraître, mai 2000).
- Baslé, (M.), « L'évaluation des politiques publiques à différents niveaux de gouvernement : moyens, portée et contraintes » in *L'évaluation : alibi ou culture d'organisation*, troisièmes assises de la fonction financière et du contrôle de gestion des collectivités territoriales, ouvrage édité par la communauté urbaine de Lille, 1999, p. 47-58.
- Baslé, (M.), « L'évaluation territorialisée des politiques publiques en France : organisation, constat et problèmes spécifiques » in *Actes du congrès de la Société française de l'évaluation*, Marseille, juin 1999, à paraître, miméo, juin 2000.
- Baslé, (M.), « International Needs for Evaluation of Regional Policies : the Case of Light Evaluation in Brittany, 1994-1996 » in *L'information régionale au service de la politique régionale en Europe*, Actes du quatrième séminaire du CEIES, Luxembourg, EUROSTAT, Commission européenne, Office des publications officielles des Communautés européennes, 1999.
- Baslé, (M.), « Contexte et méthodes pour l'évaluation partenariale des politiques territoriales en France : quelques leçons de l'évaluation des contrats

de Plan État-région et du cas breton », in *Revue d'économie régionale et urbaine*, n°1, 1998 .

Baslé, (M.) et Pelé, (F.), « Étude de cas d'une évaluation d'une politique publique européenne d'aide au développement du tourisme rural », in *Revue d'économie régionale et urbaine*, n° 4, 1994.

Baslé, (M.) et Pelé, (F.), « L'évaluation des politiques publiques et des contrats de Plan État-région en France : la situation à fin 1994 », *Cahiers Économiques de Bretagne*, n° 4, 1994.

Breton, (Y.), *Competitive Governments : An Economic Theory of Politics and Public Finance*, Cambridge University Press, 1996.

Buchanan, (J.), et Congleton, (R.D.), *Politics by Principle, not by Interest. Toward Nondiscriminatory Democracy*, Cambridge University Press, 1998.

Cioffi-Revilla, (C.), *Politics and Uncertainty. Theory, Models and Application*, Cambridge University Press, 1998.

Comité régional d'évaluation en Bretagne, *Évaluation légère du contrat de Plan État-région de Bretagne, 1994-1999*, miméo, 7, rue du général Guillaudot, 35069 Rennes cedex, 1999.

Comité régional d'évaluation en Bretagne, *Le dispositif régional d'évaluation des politiques publiques de la région Bretagne et ses réalisations*, rapport établi par Maurice Baslé et Franck Pelé, décembre 1999, CRE, 7, rue du général Guillaudot, 35069 Rennes cedex.

Commission européenne, *Évaluer les programmes socio-économiques. Solutions techniques pour évaluer dans un cadre de partenariat*, Collections MEANS, volume 4, 1999.

Conseil scientifique de l'évaluation, *Petit guide de l'évaluation des politiques publiques*, ouvrage collectif, La Documentation française , 1996.

Cour des comptes, « L'État et les contrats de Plan État-région » in *Rapport public* 1998, Éditions des Journaux Officiels, 1999.

Greffe, (X.), *Gestion publique*, chapitre 10, « L'évaluation des actions publiques », Dalloz; 1999.

Croizat, (M.), *Le fédéralisme dans les démocraties européennes*, collection Clefs Politique, Montchrestien, 1999.

Du Granrut, (C.), La citoyenneté européenne. Une application du principe de subsidiarité, *LGDJ*, 1997.

Guihery, (L.), Fédéralisme fiscal et redistribution. Fondements et enseignements du fédéralisme allemand, *thèse pour le doctorat de sciences économiques, université Louis-Lumière Lyon-2*, 1997.

Isaia, (H.), L'évaluation des nouveaux contrats de Plan État-région, *LGDJ*, 1999.

Muller, (P.), et Surel, (Y.), *L'analyse des politiques publiques*, Collection Clefs Politique, Montchrestien, 1998.

- Oates, (W.), *Fiscal Federalism*, Harcourt Brace Jovanourtech, 1972.
- Owen, (J.M.), *Program Evaluation Forms and Approaches*, Allemand et Unwin, 1993.
- Patton, (M.Q.), *Utilization-Focused Evaluation : the New Century Text*, Sage, 1997.
- Rieper, (O.) et Toulemonde, (J.), *Politics and Practices in Intergovernmental Evaluation*, Transaction Publishers, 1997.
- Robson, (C.), *Small-Scale Evaluation*, Sage Publications, 2000.
- Société française de l'évaluation, Bulletin numéro 1, disponible au siège social de cette société , 1999.
- Toulemonde, (J.) et Fontaine, (C.), Laudren, (E.) et Vincke, (P.), « Evaluation in Partnership. Practical Suggestions for Improving their Quality », in *Evaluation*, volume 4, n° 2, 1998, p. 171-188.

Contribution III

Éléments de réflexion issus de l'expérience internationale

Par Sylvie Trosa ¹

Il est intéressant de noter un renouveau notable de l'évaluation au sein des pays de l'OCDE. La mise au point par l'OCDE d'un guide des meilleures pratiques en matière d'évaluation en est un signe notable : jusqu'alors l'évaluation était considérée comme un sujet tabou et reléguée à l'arrière-plan face à la mesure des performances car elle met nécessairement en question la pertinence des objectifs et des mécanismes de prise de décision publique. L'aspect politique de l'évaluation l'avait écarté de l'agenda de l'OCDE et de la Banque mondiale. Il est intéressant de noter qu'aujourd'hui ce n'est plus le cas.

Les principaux enjeux liés à l'évaluation dans les pays de l'OCDE apparaissent aujourd'hui plus sur le plan de la stratégie que des méthodes ², ils portent sur quand et à quel niveau l'évaluation doit être organisée, effectuée et managée pour avoir le maximum d'effets sur la décision publique. La question des méthodes n'est pas reléguée au second plan mais elle apparaît plus connue et explorée que celle des stratégies d'évaluation.

L'évaluation devient un enjeu de la culture de la performance, plus qu'un simple outil au service de la performance. En d'autres termes son facteur clé de succès principal repose sur une culture de l'évaluation diffusée et acceptée par les fonctionnaires plus que sur des dispositifs techniques ad hoc et séparés des mouvements de modernisation. Ou tant que les fonctionnaires et les politiques ne croiront pas en l'évaluation, l'évaluation demeurera marginale.

(1) Sylvie Trosa a été rapporteure générale jusqu'en mai 2000.

(2) Ces conclusions se basent sur l'expérience de divers pays de l'OCDE : Etats-Unis, Canada, Royaume-Uni, Australie, Nouvelle-Zélande, Danemark, Suède, Finlande, CEE telle que nous avons pu l'analyser dans les documents les concernant.

À ce titre un des principaux problèmes constatés par les pays de l'OCDE est de créer des incitations efficaces pour que cette culture de la performance se généralise, tant du côté de l'offre (les maîtres d'ouvrage) que de la demande (fonctionnaires ou usagers ou parlementaires).

1. Les limites de l'évaluation dans les années 1970-1980 : l'absence d'interministérialité, le blâme et la faible légitimité de l'État

1.1. De la mesure des performances à l'évaluation

Un changement fondamental s'est produit entre les années 1970-1980 et 1990-2000. L'évaluation a trouvé sa place face à la mesure des performances.

Certains pays tels la Nouvelle-Zélande sont partis de l'idée que l'essor de l'évaluation se ferait spontanément comme un développement naturel de l'esprit de mesure et ont découvert que sans une impulsion centrale forte l'évaluation ne se produirait pas car elle pose des questions plus difficiles et complexes que la mesure.

La raison en est qu'une véritable évaluation se penche sur les impacts c'est-à-dire les résultats effectifs sur les citoyens et que ces résultats sont le plus souvent au croisement de l'action de plusieurs administrations (exemple : la délinquance ne peut être analysée sans l'ensemble du dispositif police/justice) dont les systèmes de mesure ne se complètent pas, car ils existent ministère par ministère.¹ L'évaluation oblige les ministères à travailler ensemble.

Seule une évaluation commanditée par le gouvernement peut amener les administrations et ministères à croiser leurs analyses et entraîner une collaboration entre administrations. Cette situation évoluera peut-être de façon positive à l'avenir et connaît des exceptions mais lorsque les sujets sont très conflictuels l'autorité du Premier ministre apparaît nécessaire.

L'enjeu devient alors de convaincre les agents que l'évaluation c'est-à-dire la mesure mais aussi l'interprétation des résultats, fait partie de la nouvelle culture de l'efficacité à laquelle ils doivent prendre part. Les incitations peuvent venir

(1) Eugen Bardach, « *Getting Agencies to Work Together* », *Brookings Institution Press*, 1998.

des usagers (Royaume-Uni), des Parlements (Australie), des Cours des comptes (Allemagne).

En fait les années 1970-1980 ont souligné deux dangers symétriques qui portent atteinte à l'idée d'évaluation.

1.2. Du contrôle et du blâme à l'évaluation

Ceci est en fait un changement historique important. L'évaluation dans les pays anglo-saxons avait d'abord, dans les années 1970-1980, été synonyme de jugement et de contrôle. Par exemple est-ce que le ministre fournissait bien les services qui avaient été demandés par le ministre ? Est-ce que le directeur d'agence avait bien rempli les obligations de son contrat de performance ? Dans un tel contexte l'évaluation apparaissait comme une menace et pouvait entraîner des réponses défensives ou auto-justificatives qui n'ont rien à voir avec une vraie évaluation. C'est pour cette raison que l'évaluation s'est de plus en plus séparée du contrôle.

1.3. L'évaluation est aidée mais non résumée par les informations venues du marché

Un deuxième facteur est entré en compte : tous les États qui ont une conception minimaliste de leur rôle n'ont montré aucun intérêt pour l'évaluation ¹. La conception minimaliste de l'État s'appuie sur une théorie (Chicago School of Economy) selon laquelle toute action humaine est mieux régulée par le marché ou l'auto-ajustement des individus entre eux que par l'État. L'évaluation n'est plus alors cruciale car dans la plupart des domaines le marché est sensé donner des signaux suffisants pour rectifier les actions et en temps plus rapide que tout dispositif d'évaluation. La seule raison pour laquelle ces États ont poursuivi des actions d'évaluation a été de savoir si leurs services sous-traités respectaient bien les conditions de prix et d'efficacité fixés dans leurs contrats.

Cette attitude d'esprit est en train de changer car ces mêmes États minimalistes se rendent compte que le marché fournit des informations mais quelquefois trop tard et ne donne dans le service public que peu d'analyses qu'attendent les usagers. Ainsi en Australie le «marché» du conseil et du placement des

(1) Christopher Hood, « *The Art of the State* », Clarendon Press, Oxford, 1999.

demandeurs d'emploi a-t-il été ouvert à la concurrence du secteur privé sous un gouvernement travailliste puis sous un gouvernement libéral. La différence est que sous le gouvernement travailliste il existait une autorité indépendante de régulation et d'évaluation dont la mission était de s'attacher à connaître les rapports de force au sein du marché, de définir des standards de qualité et d'alerter le gouvernement en temps réel sur les problèmes. Cette autorité a été supprimée et les résultats de l'appel d'offres de l'an 2000, à la surprise de tous, n'attribuent que 10 % du marché au ministère du Travail (au lieu de 100 % avant la réforme) ce qui oblige le ministre à se poser la question d'un plan social pour 20 000 fonctionnaires. Cet « effet de surprise » venant du marché aurait probablement pu être évité si l'autorité de régulation et d'évaluation avait continué à alimenter le gouvernement en informations, et de la sorte il ne se serait pas retrouvé dans une situation si difficile.

2. Approche scientifique et approche participative

La Société européenne d'évaluation s'interroge actuellement sur le développement des évaluations participatives (dont les fondements théoriques se trouvent dans l'ouvrage de « Fourth Generation Evaluations », Guba and Lincoln ; la Société européenne d'évaluation a son prochain congrès sur le thème de « Amener l'évaluation aux usagers : entre la société civile, le management public et le politique ») incluant les citoyens et des groupes à la base face à des évaluations qui apparaissaient plus comme des études de nature scientifique. Les évaluations de type participatif se sont développées considérablement ces dernières années dans les pays européens. Elles sont souvent qualifiées de « *empowerment evaluation* » i.e. d'évaluations donnant le pouvoir à leurs participants et renforçant leur capacité d'action dans le champ de la proposition et de l'amélioration des politiques publiques. Elles partent de l'idée que :

- une évaluation participative aura une plus grande validité car elle prendra en compte les analyses, concepts et valeurs de toutes les parties prenantes à une politique publique (Palumbo et Hallett, 1993, p. 22).
- une évaluation participative a plus de chances d'être utilisée parce que les participants se seront appropriés ses résultats (Patton, p. 60).

- une évaluation participative aurait une valeur éthique et démocratique supérieure car toutes les parties prenantes ont un accès égal à la parole et à l'analyse (Fraser, 1989, p. 311 et 312).

Néanmoins ces différents points n'apparaissent vrais que dans certaines circonstances spécifiques où en particulier :

- le nombre de parties prenantes est limité,
- un certain consensus existe sur les objectifs à atteindre,
- des règles du jeu d'égalité des parties prenantes sont acceptées.

2.1. L'approche scientifique : un instrument au service de remises en cause radicales

Par contre il se peut que dans certaines évaluations les parties prenantes soient trop nombreuses pour être toutes associées au même titre à l'évaluation. De plus certains thèmes sont portés par des oppositions de valeur et d'idéologie qui peuvent être irréconciliables et bloquer le processus d'évaluation (drogue, école etc.). En ce cas il apparaît nécessaire de disposer d'études objectives proposant des données chiffrées et des travaux qualitatifs qui permettent de canaliser les débats et de dépasser leur dimension idéologique, voire passionnelle.

Enfin, cas de figure opposé au précédent, le prix d'un accord généralisé des acteurs de l'évaluation peut être la non remise en cause de la politique et un consensus sur son maintien. En ce cas l'objectivité scientifique des études permet de dépasser soit des clivages de valeurs soit à l'inverse de faux consensus.

2.2. L'approche participative : un outil au service de l'amélioration de la mise en œuvre des politiques publiques

L'approche participative s'applique plus aux évaluations qui portent sur l'amélioration du « comment faire » tandis que l'approche dite scientifique est plus appropriée lorsqu'il s'agit de s'interroger sur la nature du « que faire et pourquoi ? ».

L'approche participative est particulièrement adaptée à toutes les circonstances où l'on trouve des acteurs qui n'ont ni les mêmes logiques ni la même culture et qu'il faut souder autour de buts et d'objectifs communs. L'approche participative est tout aussi nécessaire pour comprendre les processus de mise en œuvre aux différents échelons territoriaux concernés par la politique.

Dans son expérience le CNE ne pense pas qu'il y ait une dichotomie entre les deux approches si elles communiquent entre elles. L'approche participative apporte la connaissance des processus réels qui se passent sur le terrain et ne correspondent pas nécessairement au dispositif réglementaire. Elle permet de surcroît de comprendre la logique des acteurs (par exemple en matière de sécurité routière des départements apparemment comparables ont des résultats très différents, pourquoi ?). L'approche scientifique est toujours nécessaire, d'autant plus que les parties prenantes de l'évaluation ont des présupposés, des valeurs et des hypothèses fortes qui ne sont pas appuyés par des études ou des chiffres. L'approche scientifique permet de canaliser le jeu des valeurs des participants autour de données objectives.

Les deux approches peuvent ainsi s'entraider en particulier dans la mesure où nombre de politiques publiques mettent en jeu différents niveaux de gouvernement dont il faut tenir compte. Elles sont néanmoins différentes et peuvent donner lieu à des conclusions opposées lorsque par exemple les participants ont certaines images ou présupposés spontanés qui n'ont jamais été confirmés par des chiffres ou des études. En ce cas un lien dialectique doit s'établir entre un travail sur la représentation des participants et les données « objectives » produites par l'évaluation, ces deux pans étant aussi importants l'un que l'autre.

La précaution à prendre est qu'il y ait une bonne coordination entre les travaux quantitatifs ou « scientifiques » et les groupes de travail participatifs, sans laquelle chaque démarche risque d'aller dans des sens différents qui ne seront que difficilement conciliables. Il se peut par exemple qu'à un moment donné les représentations des acteurs se cristallisent sur des a priori qui n'ont jamais été vérifiés scientifiquement (ainsi chacun pense qu'il y a un lien certain entre la toxicomanie et la délinquance). Par contre si à l'inverse les résultats des recherches sont soumis à analyse et débat avec les acteurs sociaux, des effets heuristiques très intéressants peuvent se produire, faisant avancer tant la recherche que le débat social. De fait les deux approches sont tout aussi utiles :

- les approches scientifiques afin de vérifier, tester, prouver,

- les approches qualitatives afin de connaître les cultures des acteurs et éventuellement pouvoir les faire évoluer.

Un certain nombre de facteurs clés de succès se dégagent de l'expérience internationale.

3. La nécessité de générer une demande effective d'évaluation au plus haut niveau

L'évaluation suppose un appui au plus haut niveau, tant des politiques que des dirigeants d'administration centrale qui s'engagent à demander et utiliser l'évaluation dans le processus de décision. Cela a pu se résumer dans certains pays par le besoin de « carottes, bâtons et sermons ». Les bâtons peuvent être de refuser le financement des politiques nouvelles ou leur reconduction si des évaluations ne sont pas produites. Les carottes consistent à obliger à ce que le financement de l'évaluation fasse partie du programme. Les sermons proviennent d'un langage clair et répété des dirigeants sur l'importance de l'évaluation. La flexibilité dans la ré-allocation des ressources peut aussi être une incitation importante. Par exemple si des administrations sont encouragées à s'auto-évaluer elles devraient pouvoir réallouer les ressources d'un programme à l'autre sans autorisation préalable.

L'appui au plus haut niveau des politiques montre bien que l'évaluation n'est pas qu'une procédure managériale mais met en jeu les équilibres politiques au sein des pays qui la pratiquent.

- Équilibre entre l'exécutif et le législatif lorsque le législatif s'empare de l'évaluation comme d'un instrument de scrutinisation et de contrôle (États-Unis, Australie, Royaume-Uni).
- Équilibre entre les différents niveaux de gouvernement lorsque des transferts de compétences aux collectivités territoriales sont assortis d'obligations d'évaluation (on a ainsi oublié en France que toutes les lois de décentralisation prévoient une obligation d'évaluation qui n'a jamais été mise en œuvre).
- Équilibre entre les différentes institutions publiques lorsque l'évaluation donne lieu à des mesures d'étalonnage entre celles-ci (écoles ou hôpitaux par exemple, ainsi du Royaume-Uni et de l'Australie ou du Danemark).

4. La nécessité de produire des attentes réalistes

L'évaluation n'est pas là pour donner une réponse à toutes les questions ou des décisions clés en main. L'évaluation est un pas vers plus de connaissance et de maîtrise et non un mode d'emploi ou un kit tout fait des décisions à prendre. Ramener les attentes des parties prenantes à de justes proportions est probablement un des enjeux les plus difficiles à gérer dans l'évaluation.

5. Le besoin d'institutionnaliser l'évaluation

Ceci soulève des problèmes controversés au niveau international. Certains pays soutiennent qu'un lien trop direct avec les décisions budgétaires ne peut fonctionner car le budget est annuel, à court terme et politiquement déterminé. D'autres pays soutiennent que si l'évaluation n'est pas ancrée dans les processus de décision les plus importants elle ne restera qu'un exercice de connaissance. A contrario il est clair que des pays qui ont considérablement développé l'évaluation **mais sans la lier à aucun processus de décision (États-Unis, Canada) ont vu ces évaluations avoir un impact limité**. Le choix du temps d'une évaluation par rapport à l'agenda politique apparaît alors décisif.

6. Où rattacher l'évaluation ?

Aux ministères des Finances ? Au Premier ministre ? Aux Cours des comptes ? À des organismes totalement indépendants ? Aux ministères eux-mêmes ? L'expérience semble montrer qu'il n'y a pas une seule bonne solution et que la **diversité des lieux et des pratiques d'évaluation est une bonne chose si les institutions communiquent et se coordonnent**. Le lieu de l'évaluation dépend aussi de son objectif, plus les questions qu'elle posera seront gênantes plus son indépendance sera importante. Encore cela dépend-il de la culture de chaque pays, certains pays n'ayant pas le même sens de la pérennité nécessaire des institutions.

7. Planifier l'évaluation apparaît comme un réquisit de base

Tant pour que l'évaluation fasse sens par rapport à l'agenda politique mais aussi pour que les meilleures équipes d'évaluateurs soient disponibles. De bonnes équipes peuvent planifier leur travail plusieurs années à l'avance et ne seront pas disponibles si elles n'ont pas connaissance d'un programme d'évaluation pluriannuel.

8. S'assurer de l'opportunité des questions évaluées est aussi décisif

La plupart des évaluations se font dans des temps et des budgets limités. Il est alors important que le dialogue avec le maître d'ouvrage ait pu cibler les questions clés qui se posent à lui et se fassent dans des délais pertinents pour sa prise de décision. Les autres questions ne sont pas exclues, mais leur durée plus longue doit être négociée avec le maître d'ouvrage.

9. La publicité des évaluations et des recommandations ne fait pas encore l'objet d'un accord au niveau international

Deux objectifs se contredisent : la volonté d'être aussi transparent que possible et de susciter un débat démocratique et les effets de dérive dans les media lorsque des critiques sont publiquement mises au jour et menacent les gouvernements. La transparence complète des évaluations n'est pas aujourd'hui un principe universellement partagé, mais qui devrait être recherché autant que possible.

10. Assurer la participation de l'ensemble des parties prenantes

Dans nombre de pays la difficulté de gérer des instances trop lourdes a poussé les évaluations à associer les parties prenantes dans des processus consultatifs mais non à en faire des acteurs directs du processus d'évaluation et des instances. Il est à cet égard nécessaire de développer une réflexion préalable sur

les modalités d'association des acteurs publics : est-elle nationale ou locale ? Se fait-elle dans les instances ou au niveau de groupes de travail, etc. ?

11. Garantir la qualité méthodologique des évaluations

Aucun processus de participation ne saurait se substituer à des études rigoureuses, objectives, fondées qui apportent faits et chiffres. C'est pour cela que des mécanismes de contrôle de la qualité tels que le Conseil national de l'évaluation en France ou le contrôle par l'Auditeur général au Canada paraissent indispensables.

Ces facteurs de succès réunis par analyse de l'expérience internationale ne se veulent que de modestes jalons vers le succès de l'évaluation.

Contribution IV

Évaluation scientifique et évaluation participative

Par Christopher Pollitt ¹

Neutralisée par les parties prenantes ? Les limites de l'évaluation partenariale

Résumé analytique

Au cours de ces cinq à dix dernières années, on a constaté que l'évaluation redevenait à la mode en Europe. Elle s'est développée dans la plupart des États de l'Union européenne et de manière considérable à la Commission européenne. Dans le cadre de cette évolution générale une sous-tendance s'est affirmée en faveur d'approches « partenariales », « pluralistes » de l'évaluation. Le présent article commente la nature de ces approches, rapporte quelques-uns des avantages qu'on leur a attribués, et indique les situations dans lesquelles l'approche partenariale est susceptible de diminuer la rigueur scientifique et l'utilité de l'évaluation. Dans certaines circonstances, la « bonne vieille » méthode d'évaluation indépendante et scientifique doit être préférée aux modèles pluralistes, participatifs, ou donnant plus d'autonomie et de pouvoir aux individus.

« Collaborer : 1°. travailler ensemble avec un autre ou d'autres sur quelque chose. 2°. (péjoratif) coopérer avec, ou aider un ennemi à l'occupation de son propre pays. (Harrap's Chambers Encyclopedic English Dictionary, 1994). »

1. Introduction

On admet largement que l'évaluation a connu une forte expansion au cours de ces dix dernières années en Europe (Pollitt, 1998). Ce que l'on a peut-être moins commenté dans le cadre de cette expansion générale, c'est le remarquable développement d'une des catégories particulières d'approche de l'évaluation – ce que nous désignerons ici par l'expression « évaluation partenariale » *collaborative evaluation* (les définitions seront analysées ci-dessous). On ne peut affirmer que les approches partenariales sont dominantes mais elles sont

(1) Christopher Pollitt appartient à l'Eramus University Rotterdam.

certainement en vogue. Ces approches revendiquent un principe essentiel : l'implication directe des parties prenantes (*stakeholders*) dans la conception, la conduite et l'utilisation ultérieure des évaluations (pour un texte de référence, cf. Patton, 1992 ; et pour un exposé plus extrémiste, cf. Guba et Lincoln, 1989). Elles sont en évidente contradiction avec une approche expérimentale plus ancienne (ou, dans un sens plus large, « scientifique »), selon laquelle l'évaluation serait un processus qu'il vaudrait mieux confier à des experts ayant un niveau d'indépendance élevé et disposant de compétences méthodologiques solides, qui aboutiront à des appréciations qui ne seront pas influencées par les arguments avancés par les acteurs ayant un intérêt évident dans les résultats de l'évaluation (cf. le plus connu, Campbell, 1969 ; pour un aperçu plus général, cf. Shadish, Cook et Leviton, 1991). Elles sont aussi en contradiction avec une tradition managériale bien établie, selon laquelle les évaluations sont considérées comme des outils avec lesquels les politiques et les autorités peuvent améliorer leurs objectifs d'efficacité, d'efficacités et de réduction des dépenses dans la gestion des projets et programmes publics.

L'évaluation partenariale baigne dans la bonne image « démocratique » et « partenariale » qu'elle se donne. On est après tout à une époque de partenariats publics/privés, où plus d'autonomie et de pouvoir sont donnés aux individus et où des tentatives sont faites pour mettre en place des collectivités autogérées. Qu'est-ce qui pourrait être plus approprié que l'évaluation partenariale, comme dernière pièce au puzzle d'un nouveau système démocratique inspirant et éclairant le nouveau processus social démocratique progressiste ? De ce point de vue, l'évaluation partenariale se substitue et est perçue comme un progrès par rapport aux méthodes antérieures : les évaluations managériales des années 1980 et les tentatives d'évaluation scientifique des années 1960 et début 1970. Les « technico-rationnels » et les partisans des méthodes managériales ont été (ou devraient être) détrônés. À l'heure actuelle, les parties prenantes démocrates sauront se faire entendre et faire appliquer cette troisième voie.

Il s'agit là d'une présentation des choses un peu caricaturale. Quoique de nombreuses évaluations partenariales aient été entreprises au cours des années 1990 et que de nombreuses pages de documents universitaires aient été publiées sur ce sujet, cette mode particulière n'a certainement pas tout balayé devant elle. Elle a été davantage en vogue dans certains secteurs (particulièrement dans les domaines sociaux, de l'enseignement et dans les services de santé) que dans d'autres et elle a coexisté plutôt que triomphé sur des approches plus répandues d'évaluation, notamment économique et managériale. Certains partisans des approches partenariales font preuve d'un sarcasme venimeux à l'encontre de leurs rivaux « technico-rationnels » :

« L'approche technique et soi-disant objective de l'évaluation permet la cooptation par les acteurs exerçant le pouvoir d'élaborer les politiques et la répartition des financements. La relation de connivence qui se développe donc entre l'évaluation comme partie intégrante de la panoplie des méthodes de contrôle de gestion et de certaines catégories d'évaluation est particulièrement dangereuse... »
(Everitt, 1966, p. 176)

Dans le présent article, je propose une critique de l'approche partenariale. Mon argument essentiel est le suivant : « à chaque tâche son outil », l'évaluation partenariale est adaptée à certaines situations, non à d'autres. En effet, se lancer dans l'évaluation partenariale dans les « mauvaises » situations peut, paradoxalement, produire des résultats inutilement négatifs. Le présent article ne prétend pas comparer la totalité des différents paradigmes d'évaluation : un seul livre n'y suffirait pas (voir, par ex. : Pawson et Tilley, 1997 ; Shadish, Cook et Leviton, 1990). Il est focalisé sur l'évaluation partenariale, qui est digne d'attention en tant que telle, car elle est de plus en plus à la mode.

2. Définition de l'évaluation partenariale

Aux paragraphes précédents j'ai présenté l'évaluation partenariale comme un ensemble d'approches. Mon intention était de montrer que des distinctions existent à l'intérieur même des différentes positions programmatiques que j'appellerai « partenariales ». La preuve peut en être fournie rapidement en répertoriant les appellations employées par les partisans des différentes approches : l'évaluation autonomisante (*empowerment evaluation*) (Fetterman *et alii*, 1996), l'évaluation de quatrième génération (Guba et Lincoln, 1989), l'évaluation critique (Everitt, 1996), l'évaluation centrée sur l'utilisation (Patton, 1997), l'évaluation pluraliste (Duran *et alii*, 1995) et l'évaluation démocratique (Floc'hlay et Plottu, 1998). En outre, ces différences internes ne sont pas toutes insignifiantes. Les partisans de l'évaluation de quatrième génération par exemple poussent leur constructivisme social au point de nier l'existence de toute réalité objective (Guba et Lincoln, 1989, p. 44) alors que d'autres partisans de la coopération ne souhaitent pas aller tout à fait aussi loin (Duran *et alii*, 1995 ; Laughlin et Broadbent, 1996). Quelques partisans de la coopération critiquent violemment les évaluations faites pour servir les intérêts des dirigeants (Everitt, 1996), alors que d'autres acceptent avec pragmatisme que le point de vue de la direction ait un rôle important à jouer, et que les intérêts des dirigeants méritent d'être pris en compte sérieusement en même temps que ceux des autres parties prenantes (Patton, 1997). Certains se

contentent simplement d'envisager les évaluateurs comme détenteurs d'une position privilégiée et unique (en raison de leur savoir-faire et de leurs connaissances spécialisées), et insistent sur le développement de codes de déontologie et autres mécanismes pour encadrer le pouvoir et l'influence des évaluateurs. Cependant que d'autres encore font tout ce qu'ils peuvent pour minorer les privilèges de l'évaluateur et considèrent que l'un des objectifs principaux de l'évaluation est de donner plus d'autonomie et de pouvoir aux groupes défavorisés, cet objectif dépassant parfois en importance la production même de rapports quels qu'ils soient, l'accord sur une série de recommandations ou qu'une décision spécifique soit prise concernant le problème posé.

Ces différences internes sont importantes. Elles incluent des différences épistémologiques fondamentales concernant la nature de la réalité (allant d'un constructivisme social pur et dur jusqu'à des présupposés plus « réalistes ») et des divergences dans les orientations politiques (d'un égalitarisme radical à un pluralisme modéré). J'essaierai en cours de route de tenir compte des différences les plus marquantes.

Toutefois, j'insisterai ici sur ce qui unit les partisans de l'évaluation partenariale plutôt que sur ce qui les divise. En effet, ils se rejoignent sur le fait que la participation active des principales parties prenantes est fondamentale pour une bonne pratique de l'évaluation. *L'étendue* exacte de la participation qu'ils préconisent est sujette à variation (par exemple, faut-il inclure/exclure la conception de l'évaluation, le choix des méthodes, la conduite des enquêtes sur le terrain et le traitement des données, la rédaction et la présentation du rapport final, etc.), mais tous s'accordent sur une participation constante, substantielle et active. La simple *consultation* ne suffit pas. Guba et Lincoln, sans doute les plus extrémistes partisans de l'évaluation partenariale, envisagent de réunir toutes les parties prenantes (ils identifient un total de 14 catégories différentes) et de les convaincre de signer un ensemble de « conditions pour une dialectique herméneutique productive » (Guba et Lincoln, 1989, p. 191-204). D'autres, qui ne partagent pas nécessairement le relativisme épistémologique poussé de Guba et Lincoln, s'accordent néanmoins sur l'importance d'un « dialogue avec les participants » :

« L'évaluation centrée sur l'utilisation encourage les parties prenantes à se joindre aux évaluateurs en tant que citoyens informés d'une société mondiale d'expérimentation » (Patton, 1996, p. 384).

« L'évaluation critique entreprend un processus de dialogue avec les utilisateurs afin de fournir des occasions pour qu'ils améliorent leur compréhension, et par là même, un contrôle renforcé sur leurs vies » (Everitt, 1996, p. 180).

« Le consensus sur le principe d'une participation accrue du public dans la prise de décision a produit une nouvelle génération de méthodes d'évaluation qui tentent d'intégrer la population au processus décisionnel, c'est le cas notamment avec l'évaluation participative et avec l'évaluation autonomisante » (Floc'hlay et Plottu, 1998, p. 262).

« Nous voulons que l'évaluation soit inclusive, qu'elle serve des intérêts multiples et non uniques, qu'elle conteste et corrige plutôt qu'elle ne renforce les déséquilibres des pouvoirs existants. Nous voulons que les évaluations encouragent le fait de se sentir partie prenante, le sentiment d'être mandaté ainsi que la responsabilité parmi les parties prenantes-citoyens, à la fois concernant les défis de nos sociétés actuelles et trouver des solutions [...] Les systèmes de mesure des performances n'ont pas la capacité suffisante de satisfaire cette conception démocratique de l'évaluation, en partie parce qu'ils reflètent de façon close et autoritaire la manière dont les programmes sociaux ont rempli leurs objectifs, au lieu d'être ouverts et fondés sur le dialogue... » (Greene, 1999, p. 169).

« En exprimant leurs opinions sur une politique soumise à évaluation, les pluralistes ont adopté des méthodes visant à inclure une représentation aussi large que possible des acteurs ou parties prenantes concernés [...] Les approches pluralistes de l'évaluation mettent en lumière le fait qu'une évaluation doit tenir compte à la fois du problème à résoudre et des caractéristiques des acteurs surgissant autour de celle-ci » (Duran *et alii*, 1995, p. 55-52).

Même si j'adhère au passage précité d'après lequel la résolution du problème est un facteur important, je soutiens que, *étant donné justement que la nature du problème à résoudre varie*, rechercher la coopération active des parties prenantes peut parfois s'avérer malencontreux. Accroître au maximum la participation des parties prenantes *dans le processus d'évaluation* peut parfois restreindre le type de questions que l'on peut traiter, retarder ou diminuer les chances qu'une décision transparente ou dont il doit être rendu compte (*accountable*) puisse être prise et brouiller les limites de l'imputabilité démocratique (*democratic accountability*). En effet, dans certaines circonstances cela peut conduire à une exploitation des parties prenantes et à une diminution de la légitimité décisionnelle. Les assertions les plus excessives faites par certains partisans de la coopération (par exemple « la participation de chaque citoyen » ; Floc'hlay et Plottu, 1998, p. 263) nuisent aux arguments avancés par leurs collègues plus modérés en faveur d'une application plus limitée de cette approche.

3. Quels sont les arguments avancés en faveur de l'évaluation partenariale ?

Un certain nombre de qualités intrinsèques sont attribuées à l'évaluation partenariale, entre autres :

- A. *L'évaluation partenariale offre davantage de chances d'obtenir une validité (externe) plus grande car elle reflète les différentes perspectives, concepts, constructions, valeurs (etc.) des parties prenantes clés de la politique, du programme, du projet.*
- B. *L'évaluation partenariale a plus de chance d'être employée, car les parties prenantes auront participé au processus d'évaluation et ressentiront davantage qu'ils sont « partie prenante » dans le rapport final ou les conclusions.*
- C. *L'évaluation partenariale est supérieure éthiquement aux autres formes d'évaluation, car elle repose sur l'acceptation juste et démocratique du droit de chaque porteur d'enjeux de pouvoir s'exprimer complètement et à égalité sur la conduite de l'évaluation.*

Ainsi, concernant l'argument A, nous constatons que Palumbo et Hallet (1993, p. 22) recommandent l'adoption de l'approche de « quatrième génération » de Guba et Lincoln au motif que :

« une approche constructiviste est compatible avec une approche à méthodes multiples. Au lieu de se servir d'une méthodologie unique, ou de se fier excessivement à la validité interne, la triangulation permet de faire émerger les multiples constructions qu'ont sur la réalité du programme les parties prenantes. »

Patton (1997, p. 60) est le partisan le plus convainquant de l'argument B :

« Les évaluateurs centrés sur l'utilisation commencent par identifier et organiser les utilisateurs principaux visés. Puis ils communiquent avec ces utilisateurs au cours du déroulement de l'évaluation afin de renforcer l'engagement à l'utilisation de ladite méthode. »

Les partisans fervents de l'évaluation partenariale semblent être plus réticents à avouer leur soi-disant supériorité éthique (argument C), mais cela se lit entre les lignes. Everitt (1996) cite en l'approuvant le critère que Fraser emploie pour distinguer les éléments et les arguments acceptables de ceux qui ne le sont pas :

« nous pouvons distinguer les bonnes interprétations des mauvaises concernant les besoins des personnes. Dire que les besoins sont une construction culturelle, interprétée discursivement, n'est pas affirmer que l'interprétation d'un besoin en vaut un autre [...] .»

« D'abord, des considérations d'ordre procédural concernant les processus sociaux par lesquels ces interprétations des besoins concurrents sont produites [...] les meilleures interprétations des besoins sont celles obtenues grâce à des processus de communication aussi proches que possible des idéaux démocratiques d'égalité et d'équité [...] des considérations liées aux conséquences se justifient aussi dans l'interprétation des besoins. Ces considérations font intervenir la comparaison entre les différentes répartitions de résultats d'interprétations concurrents [...] des considérations liées aux conséquences exigent que, toute chose égale par ailleurs, les meilleures interprétations des besoins soient celles qui ne désavantagent pas tel groupe de personnes vis-à-vis de tel autre » (Fraser, 1989, p. 311-312).

Floc'hlay et Plottu (1998, p. 275) expriment la question plus succinctement :

« La participation accrue des citoyens à des étapes situées en amont du processus de décision renforce, par nature, la légitimité de l'action publique et son acceptation par la population. »

4. La critique de l'évaluation partenariale

Nous allons analyser en détail chacun de ces trois avantages supposés de l'évaluation partenariale. Nous porterons une attention particulière sur le premier argument (une plus grande validité et légitimité externe) car il semble être celui avancé le plus souvent :

A. *Une plus grande validité externe.* La réalisation de ce prétendu avantage dépend certainement d'une série de facteurs. D'abord, il y a la question pratique de savoir si l'évaluateur a la possibilité d'obtenir et de représenter avec précision les interprétations de la réalité de *toutes* les parties prenantes, et de le faire d'une manière équilibrée. Une telle procédure, si parfaite soit-elle, peut être impossible à mettre en œuvre pour de nombreuses raisons, notamment concernant des programmes complexes comportant de nombreux parties prenantes. Le temps peut manquer pour contacter les acteurs et rassembler tous les points de vue. Certaines personnes contactées peuvent refuser de répondre ou demeurer distantes et hostiles envers le processus dans son ensemble. Même si les parties prenantes participent, ils peuvent ne « jouer que pour la galerie » ou se révéler avoir des opinions si divergentes qu'elles ne pourraient pas être représentées dans le cadre d'un consensus unique et négocié. Pawson et Tilley

en donnant un exemple comique (comme à leur habitude) : « On a du mal à imaginer la mise au point d'une "interprétation commune" des revendications et préoccupations des groupes de vigilance de voisinage et des cambrioleurs étant donné qu'on part de présupposés radicalement différents sur la légitimité de cette manière de gagner sa vie » (Pawson et Tilley, 1997, p. 20). Dans ces hypothèses, les évaluateurs « orthodoxes » de la « quatrième génération » finiront par se disputer ou rendront un rapport qui ne ferait guère plus que de poser côte à côte deux conceptions du monde totalement incompatibles.

La question empirique qui se pose est de savoir avec quelles fréquences et dans quelles circonstances on peut regrouper les cadres de référence de parties prenantes multiples dans une analyse unique et intégrée d'une politique, d'un programme ou d'un projet. L'une des circonstances les plus difficiles est celle dans laquelle la décision à prendre apparaît aux parties prenantes comme se traduisant par une somme nulle. C'est-à-dire que le gain d'un porteur d'enjeux est la perte d'un autre. Les gouvernements sont souvent confrontés à ce genre de dilemmes, en particulier lors de l'élaboration du budget lorsqu'ils répartissent un volume de ressources fixes sur un nombre de programmes concurrents. Les ressources ou les services administratifs ne peuvent pas toujours être distribués à parts égales pour tout le monde – certaines choses « font masse » –, comme les gros contrats, les grands projets d'infrastructure, les constructions/fermetures d'hôpitaux ou les décisions sur la dépénalisation/pénalisation de l'avortement, de l'homosexualité, de l'usage du cannabis ou de l'appartenance à l'Armée républicaine irlandaise. Dans ces situations les parties prenantes peuvent n'avoir que très peu de raisons de faire des compromis et de parvenir à un consensus, comme nous venons de le constater récemment en Irlande du Nord sur la question du désarmement.

Les partisans d'approches partenariales ont admis ce problème au moins partiellement. Une de leurs réponses s'intitule la méthode multi-critères d'aide à la décision (*Multicriteria Decision Aid Method*) (Floc'hlay et Plottu, 1998) ou analyse multi-critères-multi-jugement (Toulemonde *et alii*, 1998). Il s'agit d'une famille de techniques conçues pour obtenir une structuration partenariale des problèmes, l'identification des critères et leur explication ainsi que la pondération des valeurs. Quoique très intéressantes, le nombre d'études disponibles sur leur application semble encore limité. Elles sont à la pointe des techniques expérimentales et nécessitent d'être maniées avec prudence et compétence. L'étude de cas effectuée par Toulemonde *et alii*, a donné une impression mitigée à l'auteur du présent article. D'une part, les techniques elles-mêmes sont astucieuses et plutôt transparentes une fois mises en œuvre, et une note pondérée a pu être calculée pour chacun des 20 projets : d'autre part, les

partenaires de l'expérience, deux organismes administratifs français – un département interministériel au centre et un conseil régional – étaient peu susceptibles d'être aussi éloignés en termes de culture, de ressources et d'attentes que les nombreuses autres parties prenantes pouvant être impliquées dans une tentative « d'évaluation démocratique » à grande échelle. En outre, ce qui était partiellement sous-jacent dans cette situation c'est que tout le monde pouvait y trouver son compte, au lieu que tous soient perdants. Une politique existante de développement urbain fondée sur une relation centre-région en constituait la toile de fond. Dans ce partenariat, chaque partie avait la possibilité d'avoir accès à des ressources plus importantes qu'elle n'aurait pu obtenir pour les mêmes besoins toute seule. Enfin, on pourrait observer que ce type de méthode peut difficilement être qualifié de « conception » de l'évaluation faite par les parties prenantes. Il s'agit plutôt d'un cas où on a demandé aux parties prenantes de fournir des valeurs pour les rentrer dans une équation pré-formatée et présentée par les évaluateurs.

Les paragraphes précédents soulèvent la question cruciale des limites de la sphère utile de l'évaluation partenariale. Il est clair que les nombreux exemples de coopération réussie cités dans la littérature sur l'évaluation constituent essentiellement des tentatives locales effectuées à une petite échelle dans le cadre de prestations fournies par des professionnels, en contact face à face (écoles, services de santé, services mise à l'épreuve et services sociaux personnels). Patton reconnaît explicitement cet aspect :

« Cette orientation est liée au fait qu'il s'agissait de petits programmes, réalisés au niveau local où l'idée de procéder à une "recherche" pouvait être intimidante... » (Patton, 1977, p. 25).

L'exemple donné par Toulemonde *et alii*, est plus substantiel (budget de 6 millions d'Écus, 1 200 projets, souvent petits, 240 opérateurs) mais il s'agissait essentiellement d'un programme de développement, c'est-à-dire de « subventions octroyées » qui ne concernait pas la vérification ou l'aspect régulateur ou l'une des autres facettes moins populaires de l'activité administrative. En outre, comme nous l'avons mentionné, les parties prenantes semblent avoir été composées de deux niveaux administratifs plutôt que de groupes plus étendus.

À l'évidence, les demandes d'évaluation débordent bien au-delà de ces contextes particuliers. Le système démocratique pourrait aussi bénéficier d'évaluations appliquées notamment au choix entre plusieurs sites d'installation de nouveaux aéroports, au rapport coût-efficacité des armes de haute

technologie, à l'impact des programmes extensifs de soutien financier en faveur d'industries spécifiques ou de certaines productions agricoles, à l'introduction de nouvelles formes de gestion financière, à la totalité de certains services de santé, à l'efficacité d'un nouveau programme national de lutte contre la criminalité organisée ou même aux avantages et inconvénients de différents modes de scrutin. Dans chacun de ces cas, les perspectives de rassembler toutes les parties prenantes principales pour une discussion approfondie sur les objectifs, les approches, la conception, les méthodes de collecte des données, la diffusion (et ainsi de suite) de l'évaluation semblent à première vue plus faibles que dans la plupart des exemples cités par Patton, Toulemonde ou Floc'hlay et Plottu.

En définitive, même si la gratification découlant d'une plus grande validité externe est partiellement ou totalement réalisée, est-il pertinent de poser la question du « coût » ? À combien de prérogatives et de techniques traditionnelles l'évaluateur doit-il renoncer dans la recherche du « trésor » que constituerait la participation ? Selon Guba et Lincoln, c'est à un grand nombre. Non seulement ces derniers nient la possibilité de généraliser les résultats d'une évaluation d'un contexte à un autre, mais ils affirment aussi que :

« Les données de l'évaluation tirées d'une enquête constructiviste n'ont ni un statut spécial ni valeur de légitimité ; elles représentent simplement une autre interprétation dont il faut tenir compte dans la voie conduisant au consensus » (Guba et Lincoln, 1989, p. 45).

D'autres rejettent ce relativisme explicitement, mais néanmoins semblent exclure une grande partie de la boîte à outils de l'évaluateur. Ainsi Kushner (1996, p. 194), par exemple, semble renoncer aux expériences, aux questionnaires et aux mesures du rendement au motif qu'ils ne sont pas « interactifs ».

Si les versions les plus constructivistes de l'évaluation partenariale sont adoptées, l'une des conséquences négatives sera que les évaluateurs ne pourront offrir d'explications causales. Pour les plus extrémistes de la « quatrième génération », ceci n'est pas une perte parce qu'ils rejettent la notion d'explication des causes et des effets dans le cadre de leur positivisme objectiviste, cela est une forme d'épistémologie autoritaire avec laquelle nous ne voulons rien avoir affaire. Mais, même chez les partisans plus modérés et moins sectaires de la coopération, il y a probablement *de facto* peu de chances pour qu'une évaluation puisse renseigner sur les causes de l'effectivité des programmes publics. Cela est dû au fait que les évaluations explicatives de ce

type nécessitent généralement, au stade initial, que des hypothèses causales aient été formulées et qu'un ou des modèles détaillés aient été élaborés, autour desquels la collecte de données et la mise à l'épreuve peuvent commencer. En d'autres termes, les évaluations explicatives exigent que les évaluateurs se concentrent sérieusement, soient précis et résolus au stade de la conception. Cependant, c'est exactement ce type d'approche qui risque d'être la cible des parties prenantes, si on leur donne un rôle important à jouer dans la conception et la conduite de l'évaluation. La probabilité que des parties prenantes très différentes arrivent à se mettre d'accord, au tout début de l'évaluation, sur un modèle précis de relations et de processus soumis à enquête, est, toutes choses égales par ailleurs, assez faible.

La validité interne doit être sacrifiée au nom de la validité externe : le premier doit-il toujours primer sur le second ? Évidemment, certains évaluateurs coopératifs reculent à ce stade. Patton, par exemple, affirme :

« L'honnêteté d'une évaluation dépend de ses choix empiriques, c'est-à-dire, son engagement à recueillir et rendre compte méthodiquement de données crédibles [...] Une approche empirique signifie la collecte de données sur les activités réelles des programmes et leurs conséquences puis la présentation de ces données d'une manière impartiale et équilibrée afin que les utilisateurs de l'information et les décideurs puissent en tirer leurs propres conclusions positives ou négatives, d'effectivité ou d'ineffectivité » (Patton, 1997, p. 358-359).

L'insistance systématique de Patton sur la méthode empirique et son hypothèque selon laquelle il existe des activités *réelles* à mesurer devrait déplaire aux constructivistes les plus radicaux. De façon plus générale, toutefois, on peut douter de l'étendue et de la fréquence avec laquelle une enquête empirique rigoureuse de l'effectivité peut être associée à une participation pure et dure des parties prenantes. Patton, évaluateur d'un savoir-faire et d'une expérience exceptionnelle, œuvrant principalement dans des cadres relativement locaux, a peut-être la possibilité de jouer ce tour de magie plus souvent que la plupart de ses collègues. Par exemple, Scriven (spécialiste de la méthodologie de l'évaluation et philosophe réputé) préconise avec insistance le maintien de la « distance » entre les évaluateurs et le personnel lié à une politique, un programme ou un projet soumis à évaluation. Il fait une distinction entre ce qu'il considère comme une « évaluation » véritable et une « évaluation conseil » ce qu'entreprennent essentiellement selon lui de nombreux autres évaluateurs, y compris Patton. Pour lui, une véritable évaluation doit inclure un effort systématique pour en préserver la validité et

exclure les causes probables de parti pris, y compris les contacts avec le personnel du programme :

« la valeur d'un expert dépend de la validité de son évaluation, et celle-ci est mise en péril par l'interaction sociale [...] la distanciation a un prix ; mais l'implication met la totalité de l'évaluation en danger » (Scriven, 1997, p. 488).

D'après lui « la soi-disant conception partenariale, sous-ensemble de la tendance à donner plus d'autonomie et de pouvoir aux individus est aussi bâclée que possible [...] » (*ibid.*, p. 486).

Traisons de cette question d'une autre manière. En 1997, la Société d'évaluation australienne a décerné un prix spécial du mérite à une étude du Bureau de l'économie des transports et des communications intitulée *Évaluation du programme du point noir* (1995). Le programme du point noir était une initiative fédérale visant la réduction du nombre et de la gravité des accidents dans les lieux où ils se produisaient le plus souvent (« points noirs »). L'avenir du programme était incertain, mais l'évaluation a contribué à en assurer la pérennité et l'a, en effet, renforcé. Cette évaluation montra que la diminution des accidents ayant causé des blessures aux points noirs servant d'échantillon et « traités » était plus de deux fois et demie supérieure à ce que l'on aurait pu attendre des statistiques d'accidents en général. Employant des techniques de coûts/avantages pour estimer le coût des accidents, l'équipe d'évaluateurs a conclu que le programme dégageait une valeur nette actuelle de 1 992 dollars australiens sur 791,8 millions et un ratio coûts/avantages de 3,9. L'analyse s'appuyait sur un ensemble complexe d'estimations économiques, d'essais statistiques et d'un modèle d'explication probabiliste des déplacements d'accidents. Il s'agissait donc, en d'autres termes, du type même d'évaluation « solide », « scientifique », « quantitative » que les partisans de la quatrième génération adoraient détester. Aucune coopération avec les conducteurs, les piétons ou les contribuables n'a eu lieu, il n'y a pas eu de longues réunions au cours desquelles les différents parties prenantes ont détaillé leur conceptualisation de la nature du problème, de ses causes et de ceux qui en étaient responsables. Cela signifie-t-il que l'évaluation du point noir était mauvaise ? Pour certains partisans de la coopération extrémiste, sans doute elle l'était, mais on ne pouvait certainement pas avancer que le rapport était inutile. En effet, il donnait exactement le type de chiffres sûrs, renforcé par des méthodes transparentes et systématiques, que les politiques et les hauts fonctionnaires pouvaient employer dans leur combat pour arracher des financements supplémentaires à un trésor public traditionnellement méfiant et réticent.

B. *Une utilisation plus grande des résultats de l'évaluation.* Il s'agit encore ici d'une question en partie empirique. La concertation avec les parties prenantes conduit-elle à une meilleure utilisation des résultats et des recommandations d'une évaluation ? De mon point de vue, cela n'est pas certain. Un grand débat a eu lieu dans le monde de l'évaluation concernant les facteurs encourageant une plus grande utilisation (et sur ce que cette utilisation pouvait bien signifier : cf., par ex., Rist, 1995 ; Weiss, 1988). « L'utilisation » semble se résumer le plus souvent à une influence sur le long terme dont l'effet est imprévisible, sur la manière dont les acteurs clés conceptualisent les questions de politique publique. Cependant, une opinion plus pertinente affirme que les responsables politiques « veulent d'abord des informations qui puissent leur être au moins d'une utilité pratique partielle » Rist 1995, p. xvii). En outre, les chances que cela arrive augmentent peut-être si les responsables politiques ont été mêlés de près à la mise au point de l'évaluation (Patton, 1997). Toutefois, ces résultats, qui sont provisoires, sont très éloignés de l'affirmation d'après laquelle les approches *démocratiques* entraînent une plus grande utilisation. Ils suggèrent que l'implication des « puissants » augmente les chances que les décisions ultérieures soient influencées par l'évaluation, mais ils passent sous silence l'objectif plus ambitieux d'implication de *toutes* les parties prenantes, puissantes ou faibles.

Il existe aussi peut-être un inconvénient à l'utilisation de certaines évaluations partenariales. Lorsqu'une évaluation est effectuée, par exemple, par le groupe d'évaluation ABC, les utilisateurs connaissent son « origine ». ABC est composé d'économistes et de sociologues (une supposition), ils sont connus pour avoir publié des travaux dans un domaine particulier en ayant eu recours à un ensemble de méthodes particulières. Leur rôle, leurs responsabilités et leurs méthodes sont transparents, donc leur conclusion et leurs recommandations peuvent être acceptées, critiquées ou rejetées. Leur rapport peut être soumis à un processus complet de discussion et de délibération démocratique. En revanche, si un rapport d'évaluation est produit grâce à un processus de consultation permanent, l'identité de ses auteurs semblera moins évidente. Aucune personne précise, aucune équipe n'assume l'entière responsabilité du texte. L'évaluation devient quasiment une sorte de processus de développement organisationnel (DO). Les éléments saillants, les comparaisons dérangeantes peuvent être éliminés à la rédaction dans l'intérêt du consensus. L'imprécision, les faux-fuyants en sont parfois les dénominateurs communs, comme tous ceux qui ont eu à rédiger en comité peuvent le confirmer. Les évaluations glissent vers la catégorie des communiqués, et en tant que telles perdent toute pertinence opérationnelle.

C. *Supériorité éthique*. Trois remarques doivent être soulevées rapidement pour la critique de l'hypothèse d'après laquelle l'évaluation partenariale serait automatiquement éthique (qu'implicitement les autres ne le seraient pas). Premièrement, le point de vue, consensuel, de la majorité sur une politique ou un programme (ou, du reste, le point de vue des déshérités opprimés) a parfois une image exécrationnelle. Dans certaines sociétés, à certaines époques, de fortes majorités ont soutenu, ou du moins toléré des politiques de persécution des minorités raciales, l'usage fréquent de la peine capitale, des taux d'emprisonnement excessifs ou des guerres qui se sont avérées meurtrières coûteuses et inutiles. Comme l'ont remarqué les théoriciens de la démocratie pendant plus de 2000 ans, l'opinion majoritaire à elle seule est loin d'être automatiquement morale, éthique ou même éclairée. Cela s'applique aussi aux « constructions » ou au point de vue des exclus et des déshérités.

Deuxièmement, même si la coopération des parties prenantes devait fournir un passeport automatique vers une bonne éthique et un discours à haute tonalité morale, il existe, comme nous l'avons vu, un nombre d'obstacles importants – pratiques et politiques – qui peuvent empêcher l'évaluateur d'atteindre un niveau de représentativité aussi élevé. Cependant, une coopération étroite ou partielle peut conduire à des recommandations d'évaluation d'un niveau éthique faible.

Ainsi, certaines parties prenantes peuvent simplement être prises en compte plus facilement que d'autres, certaines peuvent même se manifester avec empressement. La conséquence en sera la sur-représentation de ceux qui ont le meilleur niveau d'instruction, les ressources plus élevées et des centres d'intérêts mieux ciblés, ou de ceux qui simplement disposent de plus de temps. Les responsables d'administrations locales qui ont déjà organisé des réunions ouvertes au public ou des projets bénévoles le savent bien : à ces réunions l'on collecte souvent tout sauf un échantillon représentatif de la communauté « telle qu'elle est à l'extérieur ». La garantie d'une véritable « représentativité », est un défi considérable qui excède souvent ce que l'on peut raisonnablement attendre de l'évaluateur, de son pouvoir et de son influence. Combien d'évaluateurs sont dans la position qu'il faut pour insister pour que des politiques, des médecins chefs des hôpitaux, des généraux, des p.d.g., des secrétaires permanents de ministères ou même des caïds du crime organisé participent en personne à des discussions (sans doute prolongées) avec d'autres parties prenantes ? Patton propose de prévoir un minimum incompressible de quatre réunions d'utilisateurs primordiaux visés d'une durée d'au moins huit heures (Patton, 1997, p. 356). Les partisans de la coopération ont beau jeu d'écrire sur la « justice d'une procédure parfaite » (Floc'hlay et Plottu, 1998, p. 263) mais

dans le monde réel une procédure imparfaite où la *plupart* des parties prenantes se trouvent autour de la table, sauf quelques résistants ou oubliés, peut peut-être en réalité *renforcer* le manque de volonté des parties exclues de vouloir accepter les conclusions ou recommandations quelles qu'elles soient.

Les tentatives infructueuses d'accroître « l'autonomie et le pouvoir des individus » peuvent donc se traduire par un renforcement du rôle de ceux qui savent s'exprimer, au détriment des parties prenantes les plus faibles et les moins bien organisées. Par exemple, certaines évaluations en « partenariat » des programmes de dépenses de l'Union européenne semblent impliquer la participation d'un grand nombre de responsables de programmes et de bénéficiaires, mais excluent les lointains contribuables européens, qui sont en fin de compte ceux qui, grâce à leurs contributions nationales et à travers les programmes de la Commission européenne, paient pour les routes de campagne irlandaises, les ports de plaisance italiens ou les programmes d'échanges de jeunes entre États membres. Ou bien encore, par expérience personnelle, je me souviens d'un hôpital du service public de la santé britannique qui tentait d'accroître l'autonomie et le pouvoir des patients en organisant des réunions ouvertes où les malades étaient invités, dans le cadre d'une évaluation plus vaste, à exprimer leurs opinions sur la façon d'améliorer les prestations. L'intention était bonne et de nombreuses invitations furent lancées. Malheureusement, à la réunion, le résultat fut que les infirmières et les administrateurs étaient plus nombreux que les quelques patients, presque tous des personnes de race blanche, ayant dépassé la cinquantaine ou plus âgées encore. Les médecins se distinguaient par leur absence, ainsi que les représentants des minorités ethniques (qui composaient une grande partie des patients de l'hôpital), ainsi que les hommes, même si l'on n'insistera pas sur ce point. Les grandes quantités de thé, de café et de gâteaux ont à peine été entamées.

On observe là un phénomène asymptotique particulier. Les absents à de nombreuses réunions participatives sont en toute probabilité à la fois les plus et les moins puissants. On peut supposer que dans de telles circonstances l'évaluateur doit résoudre par obligation déontologique le casse-tête méthodologique consistant à « faire parler » ceux qui sont privés de droits et « ceux qui sont trop puissants pour perdre leur temps ». La plupart des textes des partisans de la coopération ne mentionnent quasiment pas l'approche à employer pour « prendre la place des absents ».

Troisièmement enfin, la coopération a peut-être bien une valeur éthique positive, mais elle ne constitue en aucun cas la seule considération éthique.

Dans certaines circonstances, la valeur de la coopération devra être contrebalancée par d'autres considérations éthiques, par exemple, les évaluateurs peuvent être contraints par des obligations de respect de la vie privée et de confidentialité de ne pas partager des informations significatives et pertinentes avec les autres parties prenantes. En outre, d'après la déclaration sur les principes directeurs de la Société américaine d'évaluation de 1994 : « Les évaluateurs ont des obligations qui doivent prendre en compte l'intérêt public et le bien public », et ces derniers peuvent parfois entrer en conflit avec les intérêts personnels exprimés par les parties prenantes (cité dans Newman et Brown, 1996, p. 208).

* * *

Bibliographie

- Everitt, (A)., « Developing Critical Evaluation », *Evaluation*, 2,;2,1996, p. 173-188.
- Fetterman, (D).; Kaftarian, (S). et Wandersman, (A). (dir.), *Empowerment Evaluation: Knowledge and Tools for Self-Assessment and Accountability*, Londres, Sage , 1996.
- Floc'hlay, (B). et Plottu, (E), « Democratic Evaluation: From Empowerment Evaluation to Public Decision-Making », *Evaluation*, 4,3, juillet 1998, p. 261-277.
- Fraser, (N.), « Talking About Needs : Interpretive Contests as Political Conflicts in Welfare State Societies », *Ethics*, 99, janvier 1989, p. 291-313.
- Greene, (J)., « The Inequality of Performance Measurements », *Evaluation*, 5,2, avril 1999, p. 160-172.
- Guba, (Y). et Lincoln, (E), *Fourth Generation Evaluation*, Londres, Sage, 1989.
- Kushner, (S), « The Limits of Constructivism in Evaluation », *Evaluation*, 2,2, avril 1996, p. 189-200.
- Laughlin, (R) et Broadbent, (J), « Redesigning Fourth Generation Evaluation. An Evaluation Model for Public Sector Reforms in the UK ? », *Evaluation*, 2,4, octobre 1996, p. 431-451.
- Newman, (D). et Brown, ®., *Applied Ethics for Program Evaluation*, Thousand Oaks, Californie, 1996, Sage.
- Palumbo, (D). et Hallett, (M)., « Conflict Versus Consensus Models in Policy Evaluation and Implementation », *Evaluation and program planning*, vol. 16, 1993, p. 11-23.
- Patton, (M.Q), *Utilisation-Focused Evaluation : The New Century Text* (3^e édition), Londres, Sage, 1997.
- Pawson, (R). et Tilley, (N), *Realistic Evaluation*, Londres, Sage, 1997.
- Pollitt, (C) « Evaluation in Europe : Boom or Bubble ? », *Evaluation*, 4,2, avril 1998, p. 214-224.

Rist, (R)., « Introduction », p. xiii-xxvi in Rist (R) (dir.), *Policy Evaluation: Linking Theory to Practice*, Aldershot, Edward Elgar, 1995.

Scriven, (M)., « Truth and Objectivity in Evaluation », p. 477-500 in Chelimsky (E) et Shadish (W) (dir.), *Evaluation for the 21st Century : a Handbook*, Thousand Oaks, Californie., Sage, 1997.

Shadish, (W); Cook, (T). et Leviton, (L)., *Foundations of Program Evaluation: Theories of Practice*, Londres, Sage.,1991.

Toulemonde, (J.), Fontaine, (C), Laudren, (E). et Vincke, (P)., « Evaluation in Partnership , Suggestions for Improving their Quality », *Evaluation*, 4,2, 1998, p. 171-188.

Weiss, (C), « Evaluation for decisions : is anybody there ? Does anybody care ? », *Evaluation Practice*, 9, 1, 1998, p. 5-19.

Contribution V

Évaluation et décision budgétaire

Par Keith Mackay ¹

Le développement du système d'évaluation australien

1. Introduction

L'administration fédérale australienne attache la plus grande importance à l'évaluation de ses programmes et à ce que les résultats en soient utilisés. La méthode employée pour développer les capacités d'évaluation a consisté à associer les exigences formelles de l'évaluation avec sa promotion énergique par un ministère central puissant (le ministère des Finances australien). Cela a permis d'intégrer l'évaluation aux décisions budgétaires et à la gestion quotidienne des programmes de l'État.

La conséquence en est une forte activité évaluative. Les évaluations sont même utilisées pour aider aux prises de décision du gouvernement, à la hiérarchisation des choix budgétaires et pour épauler le management des programmes dans les services opérationnels.

Chaque pays est unique et la réussite australienne est obtenue grâce à une infrastructure publique solide, en cours de développement, ainsi qu'à un niveau élevé de capacités institutionnelles et humaines. Néanmoins, les multiples enseignements de l'expérience australienne sont suffisamment pertinents pour contribuer à accroître les capacités d'évaluation dans d'autres pays.

Le présent article identifie de nombreux facteurs de réussite. Ce sont entre autres :

- la création d'une stratégie d'évaluation appliquée à l'administration tout entière ;

(1) Keith Mackay a été durant dix ans responsable de la politique d'évaluation en Australie.

- l'existence d'un ministère central puissant, partisan convaincu de l'évaluation, et qui a en permanence trouvé et suscité de nouvelles occasions pour accroître l'influence et le développement de l'évaluation ;
- un engagement et un soutien sur une décennie de la stratégie d'évaluation ;
- la mise en œuvre de réformes connexes dans le secteur du management public qui ont donné une autonomie considérable aux managers opérationnels et qui met l'accent sur les résultats nets ; ces réformes ont servi de mécanisme d'incitation pour mener des évaluations et en utiliser les conclusions.

2. Origine et étapes du développement

2.1 L'origine des réformes du secteur public remonte à 1983

L'élection d'un gouvernement travailliste réformiste en 1983 a créé un contexte favorable à l'évaluation. Le nouveau gouvernement était déterminé à améliorer l'efficacité du secteur public tout en comprimant les dépenses publiques grâce à la procédure budgétaire annuelle.

Une série de réformes dans le management du secteur public a été mise en œuvre au cours des premières années du nouveau gouvernement. Un aspect incitatif de ces réformes est la prise au sérieux du fait que les fonctionnaires sont majeurs, c'est-à-dire que « les managers gèrent » par délégation de pouvoirs et de responsabilités, traduction d'une philosophie énonçant que les managers du secteur public seront fortement encouragés à améliorer leurs performances si on leur octroie une autonomie plus grande et la capacité de gérer leurs services avec moins de contrôles exercés par les ministères centraux et moins d'ingérences. L'autre aspect qui relève plus de la sanction a été de « forcer les cadres à gérer », et cette menace est directement liée à de fortes pressions budgétaires (1,25 % de prélèvement de productivité par an pour chaque ministère).

Les changements réels comprennent notamment :

- une autonomie substantielle accordée aux ministères pour leurs coûts de fonctionnement (y compris les traitements), mais ces frais administratifs sont strictement plafonnés en valeur nominale ;

- pour l'avenir, une plus grande sécurité concernant la disponibilité de ressources pour les managers des services grâce à un système de prévisions budgétaires sur trois ans des dépenses administratives et de celles de tous les autres programmes ;
- une réduction majeure par fusion du nombre de ministères (qui depuis 1974 se limite à 18), afin de moins disperser les conseils politiques et afin d'encourager une redistribution des ressources internes grâce à une budgétisation par grand ministère.

Un ensemble de principes connexes a été inscrit dans le Programme d'amélioration du management financier, couvrant le management de programme et la budgétisation (équivalent de la RCB). Ces principes insistaient sur l'importance pour les managers des services de s'assurer que les objectifs de programme étaient réalistes, afin d'assister et de guider les managers et le personnel. Ces principes traitaient tout particulièrement de l'efficacité et de l'efficacités des programmes – la performance des programmes – grâce aux bonnes pratiques du management, la collecte d'informations sur la performance et la pratique périodique de l'évaluation. Un guide imprimé faisant l'éloge des avantages attachés à ces principes en expliquant clairement les concepts, a été publié par le ministère des Finances (MdF) et par celui qui s'appelait à l'époque le *Public Service Board*, (Commission de la fonction publique) un autre ministère central.

Les mesures de réforme du secteur public furent introduites dans un climat de rigueur budgétaire. Des considérations macro-économiques ont été une motivation suffisante pour inciter le gouvernement à réduire la part fédérale des dépenses en PNB de 30 % en 1984-1985 à un peu plus de 25 % en 1989-1990. L'impact de ces réductions de programmes publics a été plus élevé que ce que pourraient suggérer ces statistiques brutes, car le gouvernement était aussi décidé à augmenter les prestations sociales pour les personnes les plus démunies. Il y parvint en ciblant les prestations de manière précise par une vérification stricte des seuils maximums d'attribution, ce qui conduisit à une réduction substantielle des prestations sociales allouées aux « classes moyennes ».

Le MdF était l'un des architectes principaux de ces réformes, qui reflétaient son rôle de coordinateur du budget et de contrôleur des dépenses des autres ministères. Il cherchait à prendre ses distances par rapport aux détails de la question des dépenses qui le paralysaient dans des conflits avec les ministères concernant des demandes de dépenses mineures. Deux aspects des dépenses budgétaires préoccupaient le MdF : la priorité à la réduction des dépenses de

l'État et la recherche des moyens de rendre les dépenses plus efficaces et efficaces.

2.2. Un intérêt croissant porté à l'évaluation

La préoccupation d'un bon rapport « qualité/prix », c'est-à-dire l'efficience et l'efficacité des dépenses publiques, a préparé le terrain pour la fourniture par le MdF de conseils aux autres ministères sur l'évaluation de leurs programmes. Le MdF a officiellement lancé l'opération par la publication d'un manuel de l'évaluation en 1986.

À peu près à la même époque, il régnait une certaine inquiétude au sein du MdF et des autres ministères centraux sur le manque de progrès réels des services opérationnels dans le management de leurs performances. Au début de 1987, le MdF a obtenu que le Conseil des ministres adopte une règle formelle exigeant que toute nouvelle proposition de politique soumise à la délibération du Conseil devait inclure une déclaration d'objectifs et des mesures de performance (évaluation ex-ante), et qu'elle devait prévoir des dispositions pour son évaluation à l'avenir. Les ministères opérationnels et leurs ministères avaient l'obligation de mettre au point des projets pour la surveillance et l'évaluation systématique et globale de la performance de leurs programmes, et de rendre compte de ces réexamens. Le MdF devait être informé des projets d'évaluation des services.

Le MdF a renforcé son comité consultatif initial, en fournissant des documents d'orientation supplémentaires en 1987 et en offrant un stage de formation élémentaire à l'évaluation. Au début de 1988 toutefois, les projets d'évaluation élaborés par les ministères pouvaient être décrits au mieux comme parcellaires (beaucoup étaient de faible niveau), et il devenait évident qu'un examen plus poussé des pratiques d'évaluation dans les ministères s'imposait. Ainsi, le MdF a lancé une étude de diagnostic pour réexaminer les progrès de l'évaluation dans les ministères et l'état global des activités d'évaluation dans le service public. L'étude conclut :

- à un manque d'intégration de l'évaluation dans les prises de décision collégiales et financières ;
- que les évaluations avaient tendances à se focaliser sur l'efficacité et les problèmes de processus plutôt que sur la question plus fondamentale de l'effectivité globale d'un programme (si les programmes remplissaient leurs objectifs) ;

- à un niveau faible de compétences en matière d'évaluation et de capacité d'analyse ;
- et que le rôle des ministères centraux, particulièrement celui du MdF, était incertain.

2.3. Une stratégie d'évaluation formelle

L'étude de diagnostic prépara le terrain pour la requête que le ministre des Finances soumit au Conseil à la fin de 1988 demandant – et obtenant – son accord pour qu'une stratégie d'évaluation formelle permanente soit appliquée à tous les ministères. Un des principes clés et toujours en vigueur sous-jacent à cette stratégie stipulait que « la responsabilité principale pour la détermination des priorités d'évaluation, la préparation des projets d'évaluation et la conduite de l'évaluation reposait sur les services opérationnels ».

C'est par une décision officielle que le Conseil a donné son accord sur la stratégie d'évaluation. Pour l'administration fédérale et ses fonctionnaires, les décisions du Conseil ont quasiment force de loi. En effet, les audits d'efficacité conduits par le Bureau national australien d'audit (ANAO) vérifient souvent si les décisions du Conseil ont été bien respectées. Ainsi, le fait que les initiatives de réforme du secteur public étaient contenues dans des décisions du Conseil a encouragé l'ANAO à jouer un rôle de surveillance.

La stratégie d'évaluation avait trois objectifs principaux. Elle encourageait les managers de programmes dans les ministères à se servir de l'évaluation pour améliorer les performances de leurs programmes. Elle a fourni des données fondamentales sur les performances des programmes aidant à la prise de décision du Conseil des ministres et à la hiérarchisation des priorités, concernant en particulier la procédure de budget annuel, lorsqu'un nombre élevé de propositions concurrentes sont défendues par chaque ministre. Enfin, la stratégie avait pour but de renforcer la responsabilité dans un environnement déconcentré en fournissant la preuve formelle de la surveillance des programmes par les managers et du management des ressources des programmes. L'accent était mis sur la transparence, qui intéresse le Parlement, en particulier le Sénat, lors de la procédure d'examen et de vote du budget.

La stratégie d'évaluation a fourni le cadre et le ressort sous-tendant les progrès de l'ECD dans l'administration australienne depuis cette époque. Ces éléments clés incluaient les quatre exigences formelles suivantes :

- que chaque programme soit évalué tous les 3 - 5 ans ;

- que chaque fonction ministérielle (comprenant un service opérationnel plus des agences déconcentrées) prépare un plan annuel d'évaluation par fonction (PEP), couvrant trois ans d'avance, et le soumette au MdF. Les PEP incluent des grandes évaluations de programmes ayant des implications sur les ressources ou les politiques importantes ;
- que les nouvelles propositions de politique des ministres contiennent une déclaration prévoyant des dispositions pour leur évaluation à l'avenir (évaluation ex-ante et projet d'évaluation concomitante) ;
- et que les rapports sur les évaluations effectuées soient en principe publiés, sauf si des considérations importantes concernant des questions politiquement sensibles, de sécurité nationale ou de secret commercial s'appliquent, et que les documents budgétaires que chaque ministère présente au Parlement annuellement mentionnent aussi les conclusions des évaluations importantes.

Un des aspects essentiels de cette stratégie d'évaluation a été le rôle que le Conseil a consenti à donner au MdF. Le Conseil a déclaré qu'il s'attendait à ce que le MdF ait l'occasion de contribuer aux plans annuels d'évaluation et au mandat des évaluations particulières pour s'assurer qu'elles se conformaient aux politiques et priorités de tout le gouvernement, et que le MdF serait disponible pour participer directement aux choix des évaluations, sous réserve de négociations entre le MdF et les services opérationnels (ou entre les ministres si un litige apparaissait). Le MdF devait aussi fournir des conseils détaillés et des manuels de méthodologie de l'évaluation, sur les systèmes d'information et sur le management et montrer l'exemple en identifiant et en diffusant les meilleures pratiques.

Le MdF préconisait, comme élément clé des réformes du secteur public, que les ministères centraux devaient prendre du recul par rapport aux détails et se départir d'une mentalité de contrôle. Alors que la direction du MdF croyait fermement à l'intérêt de l'évaluation, elle avait soutenu que les Finances devaient être directement impliquées dans les détails, après avoir hésité cependant, étant donné que cette approche trahissait leur militantisme et leur souci d'effectuer un suivi rapproché. Le MdF ne s'est résolu à demander au Conseil qu'il lui donne un rôle important à jouer qu'après que les services opérationnels aient montré de façon claire et nette, preuves à l'appui, qu'ils ne traduiraient pas leurs belles paroles en actes en matière de management, de programme et de budgétisation, auxquels ces derniers s'étaient pourtant ralliés aussi. La stratégie d'évaluation devait s'assurer d'abord de l'existence de bonnes évaluations sur des programmes appropriés, plutôt que de créer une

culture centrée sur la performance. L'insistance sur ce dernier point n'est apparue comme la principale question importante que depuis peu de temps.

Alors que la stratégie d'évaluation n'en était qu'au stade de la proposition, la plupart des services opérationnels ont déclaré qu'ils acceptaient l'intérêt de l'évaluation mais ont exprimé leurs craintes réelles au sujet de l'implication d'un MdF envahissant. Cette adhésion déclarée, en contradiction avec la réalité quotidienne du management dans la plupart des services opérationnels, représentait une caractéristique commune à d'autres initiatives introduites quelques années plus tard concernant les mesures de performance et de management des performances.

2.4. Évolutions ultérieures

Les prochaines étapes marquantes de l'ECD en Australie ont été la publication de deux rapports, l'un émanant d'une commission parlementaire en 1990 et l'autre de l'ANAO en 1991. Ces rapports ont reconnu qu'un travail important avait été consacré à la préparation et à la conduite de l'évaluation, mais ont plaidé en faveur d'un redoublement des efforts. Ils ont notamment souligné les variations dans l'étendue des activités d'évaluation dans les différents ministères. Ils ont critiqué certains ministères pour avoir mal choisi les programmes, les thèmes principaux et les sujets traités par leurs évaluations, en particulier pour n'avoir pas insisté sur les questions d'effectivité. Ils ont soutenu que le MdF devrait encourager plus vivement les ministères à préparer et effectuer des évaluations.

À la suite de ces rapports une cellule distincte du MdF a été créée en 1991, chargée de donner des conseils sur l'évaluation, de fournir un soutien, des formations et des encouragements aux autres ministères (et aussi à l'intérieur même du MdF). Cette instance, composée de neuf évaluateurs, agissait comme le point de mire et le catalyseur de l'évaluation pour le service public australien tout entier. Cette instance partageait des bureaux avec deux autres instances, chargées de la coordination générale et du management du processus budgétaire annuel du gouvernement et de la modernisation du secteur public en général.

L'évaluation dans l'administration australienne – mesurée par le nombre d'évaluations planifiées, conduites et utilisées – était dans un état de santé florissant dès la moitié de la décennie 1990.

Toutefois, dès cette époque, le MdF était concerné par le peu de progrès des ministères à formuler des objectifs clairs et réalisables pour leurs programmes,

et à collecter et rendre compte de données valables sur les performances. Ces préoccupations ont été confirmées par deux audits commandés par le MdF concernant les rapports annuels et les documents budgétaires des ministères. Cela peut sembler paradoxal, car l'évaluation se situe en haut de l'échelle des difficultés des mesures de performance, et en général elle était effectuée correctement, cependant la détermination des objectifs de programme et la collecte fréquente d'informations sur les performances, située sur l'échelle inférieure des difficultés, étaient mal effectuées. Car en réalité, la situation reflétait l'importance prise et le succès remporté par la stratégie d'évaluation dans son effort pour encourager et prescrire l'évaluation, alors que l'on avait bien moins insisté sur le suivi des performances en cours.

Ainsi, en 1995 le MdF a obtenu l'autorisation du Conseil des ministres d'effectuer une série d'audits approfondis (échelonnés sur trois ans) des objectifs des programmes et des informations sur les performances de tous les programmes dans tous les ministères. Ces audits ont été dirigés conjointement par le MdF et chaque service opérationnel, les résultats étant remis à leurs ministres respectifs et au Conseil. Les audits s'intéressaient en partie à l'analyse de la situation existante, mais attachaient plus d'importance à comprendre comment les objectifs et les données sur les performances pouvaient être améliorés, à concevoir et à s'engager sur un plan d'action pour que ces améliorations soient réalisées.

Cela montre que les initiatives sur le management du secteur public ne nécessitent souvent qu'un coup de pouce soit donné du centre pour qu'elles soient mises en œuvre dans nombre de ministères.

L'élection en 1996 d'un gouvernement issu du *Liberal/National Party* a mis l'accent sur une diminution de la bureaucratie, une réduction des contraintes administratives et de l'obligation formelle de rédiger des comptes rendus. Le nouveau gouvernement a demandé que les programmes en cours soient réexaminés en profondeur par les ministres et leurs services ministériels afin de déterminer s'il fallait les préserver, les supprimer ou les déconcentrer. Pour les programmes qui devaient continuer à être administrés au niveau fédéral, le Conseil s'attendait à ce qu'un processus d'appel à la concurrence et de contractualisation soit mis en œuvre chaque fois que possible.

Une question soulevée par les exigences d'évaluations formelles au cours de cette période concernait leur souci de respecter les processus bureaucratiques alors que cela ne s'avérait plus nécessaire : la longueur de certaines évaluations de projets de fonctions ministérielles, par exemple, était passée d'une taille

recommandée de 20 à 30 pages à plus de 120 pages. Un consensus s'était dégagé dans la bureaucratie d'après lequel, même si les conclusions d'une évaluation étaient importantes pour aider à la prise de décision des managers de programmes et du Conseil, des projets détaillés et rédigés dans un style châtié n'étaient pas nécessaires pour atteindre ce résultat.

Un courant de pensée similaire et largement partagé affirme que les ministères ne doivent pas être gênés par des contrôles excessifs sur leurs activités internes, tant que les chefs de service et les hauts fonctionnaires sont responsables des performances, et que cette responsabilité est précisée dans leur contrat de travail. C'est une philosophie qui se contente de prendre au sérieux le fait que les fonctionnaires sont majeurs (*let the managers manage*), semblable à celle adoptée au début des années 1980. À la différence toutefois que les ministres sont plus attentifs aux performances de leurs chefs de service – comme le montre le recours accru aux contrats de performance – sans compter l'empressement manifeste des ministres à les licencier lorsque leurs performances sont jugées insuffisantes. En outre, la différence sans doute la plus importante est le progrès accompli, entre temps, dans l'implémentation d'une culture de l'évaluation, et dans l'établissement d'infrastructures dans les services ministériels pour aider à la mesure des performances et au management des performances.

Cela a conduit fin 1997 à de nouvelles évolutions de la stratégie d'évaluation : elle s'est transformée en un cadre fondé sur des **principes**, en particulier de management des performances mais non sur des obligations réglementaires. Cette démarche a été entérinée par le Conseil, et fait maintenant partie de la politique du gouvernement. Approuvée par le Conseil, une telle approche, fondée sur des principes, sert à guider les responsables des services opérationnels en mettant l'accent sur les bonnes pratiques caractéristiques du management des performances et des mesures (celles-ci comprennent l'évaluation et le suivi continu des performances). Elle traduit une forte attente en fonction de laquelle les directeurs généraux et les hauts fonctionnaires continueront à préparer, conduire et à se servir des évaluations, et elle pronostique implicitement que les progrès accomplis sont présumés acquis au jour donné. Ces questions s'adressent à la racine même du problème de savoir si les contrôles centraux doivent être stricts ou relâchés.

3. Le développement de la capacité d'évaluation : les faits

3.1. Préparation de l'évaluation

Tous les services administratifs ont préparé des projets d'évaluation de leurs ministères depuis 1987-1988. L'intention à l'origine est que ces projets devaient inclure les évaluations les plus importantes dans chaque ministère et ses agences extérieures (ces dernières années environ 160 évaluations de ce type ont été mises en œuvre à un moment donné).

Les services opérationnels eux-mêmes ont décidé quels programmes devaient être inclus dans les projets d'évaluation, et les questions que devait contenir le mandat de l'évaluation. Toutefois, le MdF tentait d'influencer généralement le choix des priorités de l'évaluation des ministères en formulant des observations directes. Ce faisant, le MdF cherchait à la fois à anticiper et à susciter les besoins d'information du Conseil des ministres. Lorsque le MdF a eu des difficultés à convaincre les ministères, il s'en est parfois référé au Conseil directement pour chercher à obtenir son appui pour une des propositions d'évaluations particulières et de son mandat détaillé ; le Conseil accepte presque toujours les propositions du MdF.

Les exigences formelles, approuvées par le Conseil, en application de la stratégie d'évaluation énonçant que des projets d'évaluation des fonctions ministérielles devaient être préparés et soumis au MdF ont certainement incité fortement les services opérationnels à élaborer des projets et à les prendre au sérieux. Un autre facteur influent a été la publication des référentiels officiels du MdF sur ce qu'il était préférable d'inclure dans ces projets, avec les mesures de suivi et les rappels faits au ministère concernant la nécessité d'élaborer ces projets. La cellule d'évaluation du MdF a mené des audits internes sur le contenu et l'étendue de ces projets d'évaluation et a fourni des retours d'information et prodigué des encouragements aux ministères et a aussi identifié des exemples de bonnes pratiques.

L'exigence formelle que tous les programmes soient évalués tous les 3 à 5 ans a aussi été déterminante dans l'établissement d'un climat dans lequel l'évaluation devenait la norme plutôt que l'exception. Le concept suivant lequel les programmes sont analysés périodiquement et de manière exhaustive encourage aussi une approche planifiée et progressive de l'évaluation. Cette exigence formelle ne doit cependant pas être prise pour argent comptant. Il est très rare qu'une évaluation unique couvre tous les aspects d'un programme. Par contre, la pratique courante est que l'évaluation soit centrée seulement sur certains problèmes clés ou certains aspects d'un programme. Le défi consiste à s'assurer

que ces questions difficiles sont réellement évaluées, c'est sur ce point que le MdF a agi par la persuasion et par son engagement direct dans des évaluations particulières.

3.2. La conduite de l'évaluation

Presque tous les ministères ont choisi de créer des cellules de coordination de leurs projets d'évaluation. Sous leurs formes la plus réduite, ces cellules se composent de deux ou trois personnes. Dans certains ministères, comme celui de l'Emploi, une cellule de 20 à 25 fonctionnaires est chargée des projets d'évaluation, de donner des conseils en matière de méthodologie de l'évaluation, de participer à des comités directeurs et de mener les évaluations les plus importantes, concernant notamment les programmes sur le marché de l'emploi.

Il n'existe pas de démarche normalisée dans les ministères sur la manière d'effectuer les évaluations, c'est une décision considérée comme étant du ressort des services opérationnels. Certaines évaluations comportent un large éventail de porteurs d'enjeux internes et externes, qui participent soit à des comités directeurs d'évaluation, soit moins souvent, aux équipes d'évaluation elles-mêmes. Certaines évaluations sont menées par une cellule d'évaluation centrale, mais plus souvent ce sont les secteurs responsables des programmes opérationnels qui s'en chargent. Ces secteurs opérationnels sont dans ce cas responsables devant les directeurs généraux du ministère sur la qualité et la rigueur de l'évaluation.

Pour les évaluations plus importantes – celles mentionnées dans les projets d'évaluation des fonctions ministérielles – une intervention externe est plus courante, par la fourniture de propositions et de commentaires sur le mandat et la méthodologie de l'évaluation proposée, la participation à des comités directeurs, la fourniture de commentaires sur les avant-projets de rapports d'évaluation.

Le grand nombre d'évaluations en cours, et le fait que plus de 530 rapports d'évaluation ont été publiés au cours des quatre dernières années environ, sont la preuve d'une vaste activité évaluative de l'administration australienne. Tout cela alimente une bibliothèque grandissante de conclusions d'évaluation.

Le MdF édite un répertoire des rapports d'évaluation publiés, cela aide à suivre l'évolution des activités de chaque ministère. Mais surtout, il aide à la diffusion des pratiques d'évaluation et des méthodes parmi les ministères ce qui fournit

une certaine assurance qualité car la disponibilité publique de ces rapports les expose à l'examen attentif des collègues. Une étude récente de la Cour des comptes australienne couvrant tous les ministères a montré que 75 % des évaluations conduites en 1995 et 1996 ont été rendues publiques ou étaient disponibles sur demande.

3.3. Qualité de l'évaluation

La qualité des rapports d'évaluation est l'aspect le plus difficile à mesurer. La rigueur d'une évaluation de programme dépend de l'expérience et de l'objectivité des évaluateurs. Une évaluation récente de la qualité sur un petit échantillon de rapports d'évaluation a été établie à la demande de l'ANAO. Il a constaté des faiblesses méthodologiques dans plus d'un tiers des échantillons. Certes, quelques évaluations publiées sont de faible qualité : certaines ont peut-être été effectuées dans un but auto-promotionnel, notamment pour justifier le maintien ou la prorogation d'un programme.

L'expérience du MdF en matière d'évaluation montre que leur qualité varie énormément. Peu importe de savoir si l'on doit s'en inquiéter, ce qu'il faut considérer ici c'est l'utilisation qui est faite de l'évaluation. Si son destinataire prévu est le Conseil des ministres (pour l'aider dans sa prise de décision) ou le Parlement (information du Parlement) alors la faible qualité d'une évaluation ou le fait qu'elle puisse induire en erreur est extrêmement inquiétant. Certes, le MdF a montré qu'il savait faire part au Conseil de ses critiques concernant la qualité d'une évaluation lorsque celle-ci était utilisée par un service opérationnel pour tenter d'influencer les débats du Conseil (les services opérationnels tenteraient normalement d'éviter de tels désaccords, que le Conseil ne manquerait certainement pas de sanctionner).

La Cour des comptes australienne a constamment soutenu que les ministères devraient mettre en place des mécanismes de surveillance de l'assurance qualité des évaluations menées par les secteurs opérationnels au sein d'un ministère. Les faits tendent à démontrer que pour les quelques ministères qui ont appliqué cette démarche, cela a été un moyen efficace de faire jouer les compétences et l'expertise nécessaire en matière d'évaluation et de s'assurer de leur qualité.

3.4. L'utilisation de l'évaluation

Le facteur décisif est la question de l'ampleur réelle d'utilisation des résultats de l'évaluation.

Une étude récente de la Cour des comptes australienne sur les services ministériels a constaté que l'impact ou l'utilisation de l'évaluation était la plus notable en ce qui concerne les améliorations de l'efficacité opérationnelle et, dans une moindre mesure, les décisions de répartition ainsi que la conception des améliorations de la qualité des services aux clients.

Une étude de 1994 sur les activités d'évaluation dans l'un des services ministériels les plus importants, celui du ministère de l'Emploi, de l'Enseignement et de la Formation (DEET), a constaté une très large utilisation des conclusions d'évaluation. L'étude, menée conjointement par le DEET et le MdF, a analysé un large échantillon d'évaluations réalisées et en a conclu que :

- 55 % d'entre elles avaient entraîné des modifications dans le management du programme ;
- 8 % d'entre elles avaient amélioré la qualité des résultats immédiats et quantifiables des programmes ;
- 10 % d'entre elles avaient eu une influence sur les nouvelles propositions de politiques transmises par les ministres opérationnels au Conseil pour délibération ;
- mais qu'il y avait une grande disparité au DEET dans la qualité et l'utilisation des conclusions d'évaluation : certains secteurs opérationnels du DEET pratiquaient une qualité élevée, alors que d'autres secteurs étaient soit invariablement faibles, soit inégaux. *Source : Crossfiled et Byrne (1994).*

Les faits attestent que les évaluations ont été largement utilisées dans le processus budgétaire. Le MdF a mené plusieurs enquêtes sur le degré d'influence des conclusions des évaluations sur les propositions budgétaires soumises au Conseil. Ces enquêtes ont été menées auprès des hauts fonctionnaires du MdF pour connaître leur opinion sur le degré d'influence réelle de l'évaluation. Bien que les résultats des enquêtes ne fournissent qu'une simple indication générale sur le degré d'influence, ils sont cependant révélateurs.

Dans le budget 1990-1991, environ 230 millions de dollars australiens (à l'époque environ 175 millions de dollars australiens) de propositions de

politiques nouvelles soumises par les ministres opérationnels, ont été estimées directement ou indirectement influencées par les conclusions d'une évaluation. Dès 1994-1995, l'année la plus récente où des estimations étaient disponibles, le chiffre s'était accru de 2,3 milliards de dollars australiens (mesurée en dollars, la proportion de nouvelles propositions de politiques influencées par l'évaluation est passée de 23 % à 77 % au cours de la période). Dans la plupart des cas, le MdF a estimé que l'influence de l'évaluation avait été directe.

L'évaluation peut aussi avoir une influence significative sur le « choix des économies » proposées par le MdF ou par les différentes fonctions ministérielles à l'examen du Conseil au cours du processus budgétaire (le choix des économies correspond à des dépenses du gouvernement qui pourraient être allégées ou entièrement supprimées). En 1994-1995 environ 500 millions de dollars australiens de choix d'économies, soit 65 % du total, ont été réalisées sous l'influence des conclusions d'évaluations.

La question qui doit être appréciée à sa juste valeur concerne les limites plausibles de l'influence de l'évaluation sur les prises de décisions des ministres ou du Conseil. Le paradigme de l'évaluation dans un projet d'investissement est typiquement celui de l'analyse des coûts-avantages : un projet se justifie si, et uniquement si, son ratio avantages-coûts est supérieur à un. Mais l'évaluation de programme est une science plus qualitative : elle peut aider à identifier l'efficacité ou l'efficacités de programmes existants ou en cours d'application mais peut rarement fournir une conclusion globale sur la valeur de l'activité. Pour prendre un exemple, si un gouvernement décide d'affecter une large proportion des dépenses aux chômeurs, les conclusions de l'évaluation du programme peuvent aider à prévoir les conséquences probables d'autres choix d'interventions sur le marché de l'emploi, tels que des subventions salariales, la création d'emplois publics ou la réglementation du marché de l'emploi. L'évaluation de programme peut servir à identifier l'efficacité-coût de chaque type d'intervention. Mais elle n'est en général pas capable d'identifier le montant total des ressources qui devraient être affectées. Il s'agit là d'une décision politique. Au plus, les évaluateurs de programmes peuvent raisonnablement et légitimement espérer que leurs conclusions influenceront un tant soit peu les prises de décisions du gouvernement et les choix prioritaires entre propositions et programmes concurrents.

3.5. Exemple d'utilisation de l'évaluation pour aider un gouvernement à supprimer et réorienter les priorités de ses programmes

Le nouveau gouvernement était déterminé à réduire et à réorienter les priorités des dépenses de l'État dans le budget 1996-1997. La primauté était donnée à des programmes concernant le marché de l'emploi et d'autres qui y étaient liés, représentant des dépenses de 3,8 milliards de dollars australiens par an (environ 2,9 milliards de dollars) .

Le ministre de l'Emploi a expliqué que la politique globale du gouvernement était destinée à subvenir aux besoins des chômeurs de longue durée et à ceux qui couraient le risque de le devenir. Cette orientation a été adoptée à la fois pour des raisons d'équité et d'efficience, cette dernière visée grâce à l'obtention d'une meilleure adéquation de l'offre et de la demande de main-d'œuvre. En même temps, le ministre souhaitait une meilleure rentabilité des investissements sur les programmes concernant le marché de l'emploi dans un environnement budgétaire serré.

Pour orienter dans le choix des politiques à appliquer on puisa abondamment dans les conclusions d'évaluations australiennes et internationales. Le ministre a souligné l'efficacité-coût relative des différents programmes concernant le marché de l'emploi. Celle-ci a été estimée grâce au calcul d'un indicateur essentiel : le coût net pour l'État de chaque embauche *supplémentaire* dans les différents programmes mesurée sur la base de la probabilité *accrue* qu'une personne aidée conserve son emploi 6 mois après avoir participé à un programme concernant le marché de l'emploi (le référentiel était constitué d'un groupe témoin de personnes similaires n'ayant pas participé à un programme).

Les résultats des évaluations indiquèrent que le programme *JobStart*, offrant des subventions salariales, a eu un coût net 4 900 dollars australiens par embauche supplémentaire, alors que le programme *JobSkills* destiné à la création directe d'emplois, a eu un coût net de 76 600 dollars australiens. Le ministre observa que le « Gouvernement va (...) concentrer tous ses efforts sur les programmes qui ont démontré leur efficacité-coût en ayant pour résultat de trouver des emplois réels. » La conséquence est que le programme *JobStart* a été conservé alors que le programme *JobSkills* a été sérieusement réduit et directement ciblé sur les chercheurs d'emplois particulièrement désavantagés.

Les économies totales dont a bénéficié l'État grâce à la diminution et la réorientation des priorités des programmes relatifs au marché de l'emploi ont été d'environ 1,5 milliard de dollars australiens sur deux ans.

Le Conseil a aussi commandé une série d'évaluations majeures sur ses nouveaux programmes relatifs au marché de l'emploi et sur les nouvelles dispositions favorisant la concurrence totale entre agences de recherche d'emplois publics et privés. *Source : sénateur Vanstone (1996) ; DEETYA (1996, 1997) ; Commonwealth d'Australie (1996).*

4. Les facteurs favorisant ou freinant la réussite : ce qui a fonctionné, les défaillances et qu'elles en sont les raisons

L'analyse qui précède a identifié un certain nombre de facteurs de réussite et de facteurs la freinant. Nous allons les examiner et faire part de quelques réflexions sur leur importance relative.

4.1. Le ministère des Finances (MdF)

Le MdF a joué un rôle moteur dans le développement de l'ECD dans l'administration australienne : cela a comporté des avantages et des inconvénients. Toutefois, l'autorité considérable dont jouit le MdF parmi les ministères et avec le Conseil des ministres lui a donné la force d'influer sur l'acceptation de l'évaluation. C'est sans aucun doute le facteur le plus important du progrès considérable de l'ECD en Australie.

Le fait pour le MdF d'être responsable du contrôle de l'évaluation constitue un avantage supplémentaire car il garantit une influence directe sur les secteurs opérationnels du MdF chargés de superviser les ministères opérationnels. Avant les réformes sur la déconcentration d'il y a quinze ans, le MdF était fortement impliqué – certains diraient enlisé – dans l'examen détaillé des activités de dépenses des ministères. L'attention récente portée à l'évaluation et à la modernisation du secteur public a contribué à créer un intérêt plus soutenu à l'intérieur de ces secteurs opérationnels pour les facteurs essentiels d'effets et d'impacts et d'un bon rapport qualité/prix ; le MdF est tout simplement un acteur bien trop important dans l'organisation administrative pour lui permettre de conserver des attitudes et des activités surannées. Toutefois, la réalisation

d'un tel changement culturel, pourtant nécessaire, au sein du MdF a pris un certain nombre d'années et a entraîné des rotations importantes de personnel.

Une focalisation accrue sur un bon rapport qualité/prix a influencé par osmose la nature et la qualité des conseils de politique que le MdF fournit au Conseil des ministres. Ces recommandations se basent de plus en plus souvent sur les conclusions d'évaluation, aidant ainsi à mieux faire connaître l'évaluation par les ministères opérationnels. La participation du MdF à certaines évaluations bien choisies, offre aussi au Conseil une garantie minimum d'assurance qualité concernant les conclusions d'évaluation sur lesquelles des propositions de politiques nouvelles pourraient être fondées.

4.2. Services opérationnels

Presque tout le travail quotidien des services opérationnels est lié à la gestion de programmes en cours. L'un des trois objectifs de la stratégie d'évaluation est d'encourager les managers des programmes à avoir recours à l'évaluation afin d'améliorer les performances de leurs programmes. Parfois, cela a posé des difficultés surprenantes.

L'utilité de l'évaluation semble évidente pour ceux qui en mesurent la contribution potentielle. Les évaluateurs avancent souvent des arguments du type « Comment pouvez-vous, les managers de programmes, améliorer ces derniers et mieux satisfaire les besoins de vos clients si vous n'évaluez pas avec soin les performances des programmes ? » Malheureusement, les appels au professionnalisme des hauts fonctionnaires n'ont que peu d'effets. La plupart des managers sont conscients des bénéfices potentiels de l'évaluation, mais ne les considèrent pas comme une de leurs priorités essentielles, particulières dans un contexte de compression sévère et de réduction des ressources.

L'expérience montre qu'avec les services opérationnels une culture forte de soutien par les ministres et les administrations centrales est nécessaire dans les cas où il est prévu d'effectuer des évaluations importantes, que des ressources doivent être allouées pour en assurer la conduite et la bonne gestion, ainsi que pour assurer la mise en œuvre des conclusions. L'engagement des directeurs généraux des services sur la réalisation des programmes d'amélioration des performances est crucial pour insuffler une culture de management fondée sur les résultats.

Depuis la dernière décennie au moins, la durée de fonction des directeurs généraux a souvent été brève. Cette rotation a eu pour conséquence que certains

services ont eu des directeurs généraux différents qui n'ont pas toujours accordé la même importance à l'évaluation et à l'effort que leur service y consacrait. Alors qu'il faut du temps pour construire une culture de l'évaluation, elle peut être réduite à presque rien très rapidement.

L'obligation de prévoir des évaluations par fonctions ministérielles a contraint les services à mettre en place une infrastructure administrative pour les établir (ce qui peut être un défaut quand l'objectif du gouvernement est de réduire les lourdeurs administratives et la bureaucratie). Dans les trois quarts des services environ, il a fallu créer des commissions, généralement présidées par un directeur général adjoint du service, organiser des réunions périodiques, étudier les programmes susceptibles d'être évalués à l'avenir, et suivre l'avancement des évaluations déjà en cours. Ce travail en lui-même engendre une dynamique bureaucratique. La plupart des services font appel à leur ministre et à son cabinet pour avoir leur avis (et leur feu vert) sur les avant-projets d'évaluation prévus.

Il est difficile de savoir avec certitude comment la « situation » de l'évaluation aurait tourné en l'absence d'un partisan aussi puissant que le MdF. Certains ministères importants, comme ceux de l'Emploi et de la Santé, auraient certainement fait un effort conséquent d'évaluation quoi qu'il en soit (l'accent mis sur l'évaluation au ministère de l'Emploi datait d'avant la stratégie officielle du gouvernement en faveur de l'évaluation). Toutefois, des conversations officieuses avec des hauts fonctionnaires au sein de ces ministères ont souligné le rôle catalyseur du MdF même dans leurs ministères. Les directeurs chargés des domaines d'évaluation les plus importants dans les services opérationnels ont généralement constaté que le MdF était un allié naturel pour les aider à convaincre dans leurs ministères certains fonctionnaires plus traditionalistes à adopter l'évaluation comme un outil de management précieux.

La plupart des ministères ont choisi de s'en remettre aux managers de programme et à leur personnel pour mener les évaluations proprement dites. Cette approche déconcentrée a aidé à « intégrer » l'évaluation en tant qu'activité fondamentale de chaque secteur opérationnel et à garantir que les évaluations aient largement recours aux savoirs et à l'expérience acquise sur le programme par ceux qui le gèrent réellement. Cela a conduit à apprécier aussi à sa juste valeur la complémentarité et parfois le fait de pouvoir substituer à l'évaluation approfondie d'un programme, la pratique plus fréquente de suivi des performances par la collecte d'informations sur les performances en cours.

L'approche déconcentrée a aussi donné aux managers de programme le sentiment d'être « partie prenante » des conclusions d'évaluation. C'est un grand avantage, en contraste saisissant avec les évaluations menées de l'extérieur, les études ou les audits de performance, car dans ces situations, le manque de connaissances sur le programme et d'engagement à mettre en œuvre les conclusions a souvent gravement nuit à l'impact de ces dernières.

Toutefois, il y a aussi des inconvénients à l'approche déconcentrée. L'un d'eux est le manque de compétences sur l'évaluation dans de nombreux secteurs responsables des programmes et l'inexpérience dans la conduite des évaluations (comme l'a suggéré une étude récente de l'ANAO sur un échantillon de rapports d'évaluation).

Certains ministères ont fait face à ce besoin de compétences et d'expérience de haut niveau en mettant en place des cellules d'évaluation centralisées offrant des conseils de méthodologie et participant à des comités directeurs d'évaluation. Le ministère de la Santé a mis en œuvre l'assurance qualité de l'évaluation dans les environnements déconcentrés en s'assurant que les compétences et les ressources suffisantes étaient à la disposition des managers de programme ainsi que l'accès aux dispositifs structurels mis en place, comme les panels techniques ou les comités d'orientation. Le ministère, comme d'autres, fait des efforts importants pour la formation de son personnel afin d'améliorer leurs connaissances en matière d'analyse et de recherche.

Un autre inconvénient de l'approche déconcentrée est que le personnel d'un programme a trop souvent le nez collé à son programme pour le voir objectivement et pour poser les questions vraiment fondamentales sur ses performances et sur la prorogation du programme. La participation extérieure d'une instance d'évaluation centrale ou de collègues d'autres programmes à des groupes de travail ou des comités de direction a été une façon de résoudre ce problème. Le MdF a souvent fait partie des participants extérieurs dans des évaluations importantes et c'est un des autres moyens pour encourager l'objectivité et la rigueur.

4.3. Autres facteurs

L'effort d'évaluation pratiqué par toute l'administration et impliquant tous les services ministériels, s'est montré utile pour développer un climat général permettant d'anticiper la conduite d'une évaluation et son utilisation et a contribué à la création d'un milieu professionnel spécialisé, particulièrement à Canberra. Cela a facilité la naissance d'un marché de l'emploi dans le domaine des compétences évaluatives, y compris sur les compétences de haut niveau dans les domaines de l'analyse des données. Sur le marché de l'emploi, on trouve de plus en plus de personnel ayant travaillé pour des instances d'évaluation, dans divers bureaux d'étude économique et à l'Agence nationale des statistiques.

Le milieu professionnel de l'évaluation se manifeste notamment au travers des réunions mensuelles du Forum de l'évaluation de Canberra. Les réunions sont organisées par un comité directeur de ministères, avec le soutien du MdF, plusieurs intervenants et discutants participent à chaque réunion sur des sujets d'actualité dans le domaine de l'évaluation. En moyenne, une centaine de personnes sont présentes tous les mois.

L'échange d'idées et de bonnes pratiques lors de ces réunions a largement encouragé les prises de contacts. Plusieurs conférences sur des sujets spécifiques et des séminaires sur des problèmes particuliers concernant l'évaluation ont été organisés par le MdF et d'autres, sur des questions comme l'évaluation des conseils de politique et la relation entre évaluation et audit.

Le rôle du Parlement n'a pas été à la hauteur des objectifs ambitieux des concepteurs de la stratégie d'évaluation, qui voyaient la responsabilité devant ce dernier comme l'un des fondements de la réforme du secteur public, y compris de la stratégie de l'évaluation. En pratique, le Parlement n'a généralement pas disposé des ressources en infrastructure nécessaires, ni du recul suffisant pour s'intéresser à l'éclairage que pouvaient apporter les conclusions d'évaluation sur les performances des programmes. Alors que le Parlement exerce un rôle de contrôle global, y compris sur les dotations annuelles, il n'a pas beaucoup approfondi les questions stratégiques de performance, préférant au lieu de cela s'intéresser aux erreurs administratives susceptibles d'embarrasser le gouvernement.

Toutefois, il y a eu quelques exceptions notables à cette indifférence et étroitesse de vue du Parlement, celles-ci concernent les commissions parlementaires enquêtant sur des questions telles que le Programme d'amélioration de la gestion financière et la qualité du service. Leur utilité a été

de mettre l'accent sur les questions de performance, telles que l'impact des programmes administratifs sur les clients finaux, sur les ministères et plus généralement sur les fonctionnaires.

5. Conclusions de l'expérience australienne

La stratégie de l'évaluation formelle et normative en Australie a été un exemple de gavage centralisé destiné à garantir la préparation, la conduite, la qualité et l'utilisation de l'évaluation. Elle reflétait la conviction que les managers ne l'appliqueraient pas si on les laissait tranquilles. Les exigences formelles et une approche englobant toute l'administration a aidé à amorcer le processus et à le faire passer réellement à la vitesse supérieure. Toutefois, ce ne sont pas les règles et exigences formelles qui ont déterminé l'ampleur des évaluations menées et l'utilisation de leurs conclusions, ce sont les engagements des individus et de leurs organisations et la façon dont ils l'envisageaient ainsi que leurs motivations.

L'adoption récente d'une approche fondée sur les principes reflète une évolution dans les dispositifs relatifs à la gestion des affaires publiques et les circonstances particulières liées à la situation australienne. Cette évolution s'est traduite par le passage de contrôles stricts sur les ministères à des contrôles souples.

L'expérience australienne offre une multitude de leçons. Cependant, même s'il est possible d'identifier les caractéristiques du système ayant permis une réussite si considérable jusqu'à alors, cela ne signifie pas que ces facteurs de succès constituent des *préalables* à la réussite.

Les quelques facteurs clés du succès ont été :

- les pressions macro-économiques qui ont entraîné des budgets serrés et une recherche prioritaire sur la manière d'obtenir un bon rapport qualité/prix ;
- un ministère puissant (MdF) prêt à être un partisan chevronné de l'évaluation, et prêt à réagir aux changements de situation et à identifier les nouvelles possibilités d'influence et de développement ;

- un engagement de l'administration envers la stratégie de l'évaluation, maintenu sur une décennie, particulièrement de son partisan le plus chevronné, le MdF ;
- la présence d'une Cour des comptes prête à inciter les ministères à porter leur attention sur l'évaluation et sur le management des performances au sens large ;
- la création d'une stratégie explicite de l'évaluation avec des exigences d'évaluation formelle ;
- une stratégie appliquée à toute l'administration aidant à réaliser et préserver une dynamique dans la capacité de développement de l'évaluation ;
- une agence budgétaire (MdF) capable d'intégrer l'évaluation à la fois au sein d'un processus budgétaire et au cœur de la modernisation du secteur public ;
- un système budgétaire qui rend les coûts transparents et souligne les avantages et inconvénients de politiques concurrentes ;
- la mise en œuvre de réformes connexes dans le management du secteur public, en particulier la budgétisation par portefeuille, qui donne une autonomie conséquente aux managers opérationnels et met l'accent sur les résultats essentiels et les *outcomes*, tout ceci a créé des incitations puissantes pour les managers ;
- le soutien du Conseil et d'un certain nombre de ministres clés et le fait qu'ils aient insisté pour avoir à leur disposition des conclusions d'évaluation pour les aider à la prise de décision ;
- la priorité donnée à l'évaluation dans plusieurs grands services opérationnels importants, ce qui a contribué à la mettre en lumière et la légitimer ;
- et l'approche de déconcentration envers l'évaluation dans les services opérationnels, ce qui a contribué à intégrer l'évaluation comme activité essentielle, en même temps que les processus internes d'assurance qualité.

Contribution VI

L'évaluation des politiques et des programmes au niveau du gouvernement fédéral canadien

Par Jean Turgeon ¹

L'étude de cette période (1930/fin des années 1970) permet de mieux comprendre les assises de cette discipline dans l'appareil gouvernemental. La seconde partie dresse le tableau de l'évolution de l'évaluation de politique et de programme depuis le moment de sa mise en place « officielle » dans la structure gouvernementale jusqu'à ses développements les plus récents.

Après cette présentation de l'état de l'évaluation au sein des institutions publiques canadiennes, nous donnerons un court aperçu de la discipline à l'extérieur de ces institutions. Bien que l'appareil gouvernemental fédéral soit un acteur central dans le développement de l'évaluation au Canada, l'étude n'est pas complète si elle n'aborde pas les « autres acteurs canadiens » qui sont de plus en plus nombreux à participer à l'évolution de la discipline au Canada. Nous entendons par « autres acteurs » les gouvernements des provinces et territoires, tous les organismes regroupant les personnes intéressées par l'évaluation telles la Société canadienne d'évaluation (SCÉ), la Fondation canadienne pour la vérification intégrée (FCVI) ainsi que les chercheurs et universitaires. En conclusion, le texte propose une réflexion sur l'avenir de l'évaluation au gouvernement du Canada.

1 Quelques définitions préalables

Depuis vingt ans, différents chercheurs du secteur de l'administration publique ² et du domaine de l'évaluation ³ encouragent les organisations publiques à produire des évaluations de leurs activités afin d'améliorer la gestion des affaires publiques de l'État. Malgré ce même message, les chercheurs divergent

(1) *Jean Turgeon est professeur à l'École nationale d'administration publique au Canada.*

(2) *Weiss, 1973, Scriven 1980 ; Rossi et Freeman 1979.*

(3) *Les courants de pensée font référence aux différents fondements de la discipline.*

quelquefois d'opinion quant aux finalités que devrait poursuivre l'évaluation et quant aux méthodes pour la réaliser. Leurs argumentations se distinguent principalement en raison du courant de pensée ¹ qu'ils privilégient (Campbell et Stanley, 1963; Chen et Rossi, 1980 ; Guba et Lincoln, 1989 ; Scriven, 1997). Chacun de ces courants apporte des notions, des caractéristiques et des finalités différentes à la discipline de l'évaluation.

L'intérêt de présenter ces distinctions n'est pas d'ouvrir le débat sur les différents courants de pensée en évaluation, mais plutôt de présenter une conception de l'évaluation tout en précisant que celle-ci n'est pas la seule possible. Pour les fins du présent texte, l'évaluation dans le secteur public est considérée comme ayant pour principal objectif celui d'appuyer les processus de gouverne et de gestion des politiques et des programmes, et cela compte tenu de l'absence de lois du marché pour juger de la pertinence, de l'efficacité et de l'efficience de ces politiques et programmes (Bégin, Joubert et Turgeon, dans C. Bégin et coll. 1999, p. 267). En relation avec cette conception de l'évaluation des politiques et des programmes publics au Canada, une distinction largement répandue dans la littérature relative à l'évaluation est retenue : l'évaluation managériale et l'évaluation de légitimation.

Rappelons que l'évaluation managériale répond à des besoins de gestion ou de management de l'appareil gouvernemental. Elle permet de vérifier si l'appareil bureaucratique « fait bien les choses ». Concrètement, l'évaluation managériale peut se définir comme l'étude de l'efficacité et de l'efficience des politiques et des programmes publics. Pour ce faire, elle compare les résultats d'une politique ou d'un programme aux règles et normes édictées par les planificateurs et les dirigeants étatiques. Ce type d'évaluation se concentre souvent sur les activités de surveillance ou de suivi d'implantation de politiques et de programmes qui regroupent un large éventail de données administratives. À la lumière des résultats obtenus par l'évaluation, les administrateurs ajustent leurs activités et leurs façons de faire afin d'améliorer l'efficacité et l'efficience des programmes dont ils ont la responsabilité.

Contrairement à l'évaluation managériale, l'évaluation de légitimation ne limite pas son analyse aux éléments qui composent le processus de gestion de la politique ou du programme. L'évaluation de légitimation s'intéresse à la raison d'être, aux impacts et aux rendements des politiques et des programmes. La question qui guide alors l'évaluation devient « Est-ce que nous réalisons les

(1) Johnson, 1992 ; *Commission royale sur la gestion financière et l'imputabilité*, 1979 ; *groupe de travail ministériel chargé de l'examen des programmes*, 1986 ; *Vérificateur général du Canada*, 1983, 1987, 1990.

bonnes choses ? ». Par ce type d'évaluation, le gouvernement désire juger de la pertinence des programmes développés, mais aussi mesurer et apprécier leurs effets, leurs répercussions ainsi que leur rendement.

Dans la suite de ce texte, nous verrons que les gouvernements canadiens, compte tenu des arrangements institutionnels d'évaluation qu'ils ont privilégiés depuis plus de vingt ans, ont largement favorisé la rédaction et la diffusion d'évaluations managériales. Toutefois, il semble que le défi de réaliser de manière soutenue des évaluations de légitimation reste à toutes fins pratiques entier.

2 L'éveil de la discipline au Canada

Au Canada, les premières évaluations prennent la forme de vérification financière ou de vérification de « convenance » et cela principalement en raison de l'adoption, dès 1931, de la « consolidation du revenu et de la vérification » au sein de l'appareil gouvernemental. Bien qu'existante au Canada depuis le début des années 1930, l'évaluation de l'activité gouvernementale commence vraiment à se développer dans les processus et les arrangements institutionnels durant les années 1960 et 1970. Cette émergence de l'évaluation comme aspect incontournable du cycle de gestion est principalement due à l'acharnement des vérificateurs généraux qui se sont succédés durant cette période. Cette opiniâtreté s'exprime particulièrement dans les rapports annuels du Vérificateur général qui deviennent plus critiques quant aux dépenses gouvernementales inappropriées (Kernaghan et Siegel, 1987). Par ce geste, le Vérificateur général popularise son statut et son rôle en attirant considérablement l'attention des médias (Segsworth, 1990a). Par cette nouvelle forme d'intervention, le Vérificateur général désire élargir son mandat et développer son soutien politique à l'extérieur du Parlement. En réponse à cette façon de faire, le gouvernement fédéral tentera sans succès de limiter les pouvoirs du Vérificateur général du Canada à la fin des années 1960.

L'évaluation apparaît donc, au début des années 1960, sous l'étiquette de la « responsabilité » (McQueen, 1992). À ses débuts, elle est déjà considérée comme un élément important du cycle de gestion, complémentaire à la planification et à la budgétisation de l'appareil gouvernemental. Simultanément aux actions du Vérificateur général, le Conseil du Trésor publie en 1966 un « Guide sur la gestion financière » (directive MI-3-66). Cette initiative fait suite aux recommandations d'une vaste commission d'enquête (Commission

Glassco) qui précise le besoin de surveiller et d'évaluer la performance des activités gouvernementales en utilisant principalement des techniques modernes de gestion (Dobell et Zaussman, 1981; Segsworth, 1990b).

En 1969, le gouvernement du Canada introduit, quelques années après son introduction aux États-Unis, une nouvelle approche de gestion financière : le PPBS (*Planning, Programming, Budgeting System*). Cette approche marque le début réel de l'évaluation des politiques sociales au Canada (Segsworth, 1990b). Pour l'administration fédérale canadienne, elle permet d'implanter au niveau central (Conseil du Trésor) un système d'allocation des ressources matérielles, humaines et financières grâce à une remise en question permanente des interventions gouvernementales. Par cet exercice, les ministères et organismes sont invités, à l'intérieur d'un mémoire de programme, à décrire les besoins auxquels répond le programme, les objectifs poursuivis, la pertinence des moyens utilisés de même que les solutions de remplacement. Bien que l'administration fédérale ait réussi à implanter une budgétisation par programme, le PPBS, lui, est abandonné quelques années plus tard. Malgré tout, l'exercice a permis de développer une nouvelle approche financière orientée davantage vers les résultats.

Durant la première partie des années 1970, les évaluations fédérales sont réalisées par un service du Conseil du Trésor (Mayne, 1992, p. 4). Cette équipe d'évaluation a un mandat « horizontal », c'est-à-dire qu'elle doit aider les ministères à effectuer les réaménagements de ressources entre les différents programmes (dans un même ministère) et entre eux. Cependant, cette façon de faire change dramatiquement à partir de 1976. En effet, dans son rapport annuel de 1976, le Vérificateur général indique qu'à toutes fins pratiques le Parlement a perdu le contrôle effectif sur les dépenses et il l'encourage à prendre des mesures énergiques pour redresser cette situation (Vérificateur général du Canada, 1976). Il recommande deux changements au sein de la structure étatique que l'on peut qualifier d'importance notable pour la fonction évaluation. Tout d'abord, il demande au gouvernement que son mandat, en tant que Vérificateur général, s'étende dorénavant sur une période de dix ans afin d'assurer davantage l'indépendance de ses fonctions face au gouvernement. Ensuite, il recommande la mise en place d'une véritable politique fédérale d'évaluation de programme. Le gouvernement répondra favorablement à ces deux recommandations.

3 Les premiers arrangements institutionnels d'évaluation (1977) : la mise en place "formelle" de l'évaluation au gouvernement fédéral

C'est ainsi qu'en 1977 le Conseil du Trésor du Canada jette les bases des premiers arrangements institutionnels en matière d'évaluation de programme. Ces arrangements sont de deux ordres : la décentralisation de l'évaluation au sein des ministères et organismes fédéraux ainsi que la création du Bureau du contrôleur général du Canada, la décentralisation de l'évaluation.

Les ministères et organismes fédéraux sont invités à mettre sur pied leur propre service d'évaluation (McQueen, 1984, p. 36). La nouvelle politique du Conseil du Trésor exige du plus haut fonctionnaire de chaque organisme qu'il mette en place une fonction évaluation et que celle-ci contribue à la prise de décisions relativement aux programmes. La fonction évaluation est donc réalisée dans chaque ministère et elle répond à leurs fins propres ¹.

4 La création du Bureau du contrôleur général du Canada

Le second élément-clé des premiers arrangements institutionnels préconisés par le Conseil du Trésor est la création, en 1978, du Bureau du contrôleur général du Canada ² (BCG). Deux responsabilités majeures échoient alors à cette nouvelle entité (Kernaghan et Siegel, 1987) :

- développer les politiques en matière de gestion financière pour les services ;

Cette documentation a comme principal objectif de permettre aux ministères et organismes de réaliser adéquatement les évaluations de leurs programmes.

Pendant cette période, le contrôleur général du Canada va exhorter sans relâche les ministères et organismes à s'intéresser aux questions d'évaluation relatives à la raison d'être, aux impacts et aux solutions de remplacement à envisager pour les programmes sous leur responsabilité. Par exemple, dès 1981, il indique :

L'évaluation d'un programme peut porter sur une vaste gamme d'aspects.

(1) *Un programme est « un ensemble d'actions limitées et surtout défini plus précisément dans ses moyens et ses objectifs opérationnels » (Perret, 1994, p. 94) .*

(2) *Selon le Vérificateur général du Canada (1999, p. 3) les fonctionnaires peuvent à l'occasion être tenus de rendre compte au Parlement de ce qu'ils ont fait au nom de leur ministre, mais ils ne répondent pas de la politique gouvernementale.*

Ceux-ci peuvent être regroupés en trois catégories (Canada, 1981a) :

- maintien de la pertinence : raison d'être du programme ;
- résultats du programme : atteinte des objectifs, retombées et conséquences ;
- rentabilité : recherche d'autres programmes plus rentables et de façons plus rentables d'exécuter le programme.

Selon McQueen (1984, p. 38), les processus d'évaluation et les produits qui ont été développés avant le début des années 1990 ont contribué de façon importante au processus décisionnel au Canada. Selon l'auteur, cette contribution est principalement observable dans le domaine de la prise de décisions relatives à la conception et à la refonte des programmes ainsi que de la réallocation des ressources entre les divers programmes d'un même ministère ou organisme.

5 La situation au début des années 1990

À l'aube des années 1990, le Bureau du contrôleur général du Canada présente l'évaluation comme étant une réussite canadienne en tant qu'outil de gestion ministérielle malgré le fait que la qualité des pratiques d'évaluation varie grandement d'un ministère à l'autre (Mayne, 1992, p. 9-10).

Cependant, à la même période, d'autres acteurs de la scène fédérale remettent en question la pertinence et l'impact de l'évaluation. Ainsi, dans un rapport datant de 1991, le Sénat du Canada juge sévèrement le rôle du Bureau du contrôleur général du Canada dans les difficultés que connaît, selon lui, la fonction évaluation de programme (Canada, comité sénatorial, 1991). Selon le comité sénatorial, l'évaluation dans l'administration fédérale se préoccupe trop de corriger les dysfonctionnements des programmes (aspect managérial) et trop peu de les resituer dans leur contexte global, en ne questionnant pas systématiquement à la fois leur raison d'être, leurs impacts et leur rendement relatifs et sociaux. Le comité sénatorial exprime aussi de sérieuses réserves sur le monopole qu'exerce à toutes fins pratiques l'exécutif et se demande s'il n'y a pas lieu de créer une fonction d'« évaluateur général ». Des chercheurs analysent cette même situation d'un angle tout à fait différent. Par exemple, Johnson (1992) explique que cette tendance qu'ont les organismes et ministères

du Canada de limiter leurs évaluations aux aspects de gestion peut davantage s'expliquer par le fait que l'évaluation est trop souvent associée, dans sa situation organisationnelle, à la vérification interne. Ainsi, en limitant la fonction de l'évaluation à un exercice de vérification interne, les responsables négligent les questions concernant le rationnel et les impacts du programme.

Dans l'année qui suit, le Vérificateur général utilise son rapport annuel pour réaliser un vibrant plaidoyer en faveur de l'évaluation de programme (Vérificateur général du Canada, 1993, chap. 8). Par la même occasion, il fait l'inventaire des réalisations du gouvernement en ce domaine et constate que les ministères et les organismes ont mis en place, depuis 1983, un système d'évaluation tout à fait pertinent pour l'étude de l'efficacité des politiques et des programmes publics (Vérificateur général du Canada, 1993, p. 231). Dans son rapport, le Vérificateur général souligne que, pour les années 1991-1992, les dépenses imputables à l'évaluation de programme atteignent 263 années/personnes. Il précise que les ministères et les organismes ont consacré, en plus de cet investissement, plus de 8,5 millions de dollars pour des services contractuels en évaluation (Vérificateur général du Canada, 1993, p. 266). Malgré ces gains quantifiables pour la discipline, le Vérificateur général demeure très critique sur l'évolution de cette fonction. Selon lui, « depuis dix ans, [l'évaluation] est loin d'avoir répondu aux attentes » et cela particulièrement au moment où l'information sur l'efficacité des activités du gouvernement serait particulièrement utile pour nos décideurs publics (Vérificateur général du Canada, 1993, p. 231).

6 1994 : de nouvelles orientations en matière d'évaluation

À partir du début des années 1990, des modèles inédits de prestation de services commencent à se développer au Canada. Ils s'appuient, entre autres, sur de nouveaux partenariats avec le secteur privé, les entreprises communautaires et les organisations non gouvernementales. Ainsi, les ministères sont davantage appelés à travailler sur des problématiques communes. Ils sont souvent co-responsables de programmes conjoints où ils partagent la responsabilité de l'atteinte de résultats (ex.: l'Initiative de lutte contre la violence familiale à laquelle participent 11 ministères fédéraux). Le gouvernement canadien cherche également de nouveaux moyens qui permettraient de passer d'une gestion axée sur les règles et les procédures à une gestion qui accorde réellement plus d'attention aux résultats. C'est dans ce contexte que le secrétariat du Conseil du

Trésor approuve, en mai 1994, de nouvelles politiques d'examen, de vérification interne et d'évaluation. Ces nouvelles politiques visent à maximiser la valeur des ressources consacrées aux activités d'examen. Prises dans leur ensemble, ces trois politiques constituent pour le gouvernement un cadre devant permettre d'aligner les activités d'examen sur les priorités du gouvernement, de renforcer les pratiques de vérification interne et d'évaluation et de multiplier les avantages de l'examen à tous les niveaux. Soulignons que, pour l'heure, ces politiques sont encore en vigueur au Canada et, comme nous le verrons, elles ne remettent pas en question le monopole de l'exécutif sur la fonction évaluation au sein de l'État canadien.

7 La politique d'examen

La politique d'examen permet au gouvernement d'obtenir les renseignements pertinents et opportuns sur ce qui fonctionne bien et sur les répercussions des politiques et des programmes pour qu'il puisse prendre des décisions et rendre des comptes. Cette politique veut mettre les cadres hiérarchiques et les gestionnaires de politiques de chacun des ministères à contribution en reconnaissant que la plus grande partie de l'examen doit être exécutée par les ministères. Pour sa part, le secrétariat du Conseil du Trésor est responsable de :

- 1) coordonner ou de diriger, suivant les instructions, un nombre limité d'examens portant sur des questions concernant plusieurs ministères et l'ensemble du gouvernement ;
- 2) veiller à ce que les engagements du gouvernement envers la réalisation de certains examens soient tenus (Canada, 1994, 2).

La politique d'évaluation

La politique d'évaluation souligne que l'évaluation doit produire des renseignements opportuns, pertinents et crédibles afin que les gestionnaires et les autres parties intéressées puissent les utiliser en toute confiance. Ces renseignements concernent les questions d'évaluation suivantes (Canada, 1994, chap. 3-2 , p. 5) :

La politique ou le programme sont-ils toujours en accord avec les priorités du ministère et répondent-ils réellement aux besoins actuels ?

Atteignent-ils leurs objectifs selon le budget établi et sans enlèvement à pouvoir déterminer les effets réels des programmes fédéraux et à en estimer le

rendement ? D'autres instances du gouvernement, tel le Conseil privé, jugent, pour leur part, que l'évaluation est une activité utile pour la société. L'argumentation développée par le Conseil privé est que l'évaluation est rentable, puisqu'elle aurait contribué à la diminution des dépenses de programme principalement pendant les années 1990. Ainsi, dans un document publié en 1996, le Bureau du Conseil privé attribue à l'examen des programmes deux grands changements en ce qui concerne les dépenses gouvernementales. L'examen des programmes aurait grandement contribué à la diminution de 12 %¹, des dépenses de programme² ainsi qu'au « recentrage » des programmes fédéraux autour de programmes sociaux clairement établis. Le Bureau du Conseil privé explique ce dernier changement par la diminution importante des dépenses ministérielles consacrées aux programmes économiques au cours des dernières années.

Par ailleurs, le recours aux partenariats, notamment par la consultation systématique de larges sections de la population canadienne, s'étend dorénavant à l'étape de la formulation des politiques et programmes publics, ce qui n'est pas sans conséquences sur les manières de concevoir et de réaliser l'évaluation. Cette nouvelle façon de faire modifie à la fois la gouvernance et le management (formulation des politiques et mise en œuvre) de l'appareil d'État. En conséquence, la population exprime de plus en plus clairement aux gestionnaires son désir qu'ils accordent davantage d'attention aux résultats à obtenir plutôt qu'à l'étude de la façon de faire les choses. De plus, elle exige d'être mieux informée sur ce qu'elle reçoit contre l'argent de ses impôts (Bureau du vérificateur général, 1998, p. 5 ; Mayne, 1999, p. 1).

Conscient des attentes du public et des différentes organisations politiques qui gravitent autour de lui, le gouvernement adapte une nouvelle fois ses façons de faire et décide de faire plus régulièrement état des progrès réalisés dans l'atteinte des résultats escomptés. Cette nouvelle approche, développée à partir de la fin de 1995, exige que les ministères fédéraux modifient leur façon de gérer et de rendre des comptes aux niveaux supérieurs³. Le secrétariat du Conseil du Trésor publie, en 1996, le premier rapport sur le rendement. Il dresse dans ce rapport, le portrait de sa propre performance comme ministère de même

(1) Concrètement, chaque ministère se fixe tout d'abord un calendrier pluriannuel d'évaluation, approuvé par la haute direction. Ensuite, il incorpore les constatations des études d'évaluation au plan opérationnel pluriannuel ainsi qu'au plan de dépenses. Finalement, les plans d'évaluation des ministères et des organismes sont présentés à titre d'information au Bureau du contrôleur général (McQueen, 1984, p. 37).

(2) Le Bureau du contrôleur général est un bureau distinct de celui du Budget.

(3) Une diminution prévue entre 1993-1994 et 1998-1999.

que de celle de 15 autres ministères (secrétariat du Conseil du Trésor, 1996). Le ministre publie également son rapport sur l'amélioration de la mesure des résultats et la responsabilisation pour la fonction publique.

À partir de l'exercice 1997-1998, le gouvernement canadien entame également la mise en place progressive des principes de la fonction de contrôleur qui doit s'étendre sur la prochaine décennie. La fonction moderne de contrôleur est une réforme de gestion qui consiste à gérer soigneusement les ressources et à prendre des décisions éclairées de manière à améliorer le service, la politique publique et les résultats. Il s'agit donc de transformer la fonction de contrôleur classique, qui est étroite et spécialisée, en une responsabilité fondamentale de tous les gestionnaires de la fonction publique. Cette nouvelle fonction de contrôleur général est fondée sur l'intégration des données financières et de l'information non financière sur le rendement, sur une approche raisonnée en matière de gestion des risques, sur des mécanismes de contrôle appropriés et sur un ensemble de valeurs et de règles d'éthique communes (Canada, secrétariat du Conseil du Trésor, 1998, 1999).

Toujours avec l'objectif d'informer le citoyen sur le rendement du gouvernement, le secrétariat du Conseil du Trésor élabore, au cours de l'année 1998-1999, différents cadres d'évaluation et de responsabilisation en collaboration avec six ministères. Cette initiative s'étendra à d'autres ministères dans les mois suivants. Cette façon de faire a pour objectif d'encourager les ministères à surveiller leur rendement et à présenter des rapports fondés sur les résultats. Le rapport gouvernemental de 1998 - une gestion axée sur les résultats - comprend 80 rapports de rendement provenant de différents ministères et présente un tableau complet des engagements en matière de résultats.

Parallèlement aux préoccupations du Conseil du Trésor, le Vérificateur général du Canada ne demeure pas en reste. Il s'intéresse toujours à l'évaluation. N'ayant toujours pas le mandat de procéder lui-même à des évaluations de légitimation, il n'en demeure pas moins que l'évolution récente de la notion de vérification intégrée lui permet de s'intéresser à des questions qui, il n'y a pas si longtemps, étaient du ressort exclusif des évaluateurs de programme. En effet, les critères actuels de vérification utilisés en vérification intégrée procurent des standards permettant la mesure et l'évaluation du succès d'un programme eu égard aux objectifs d'efficacité et d'économie. En ce sens, la vérification intégrée représente une forme d'évaluation ex-post des performances managériales des responsables du programme à l'étude. Cependant, selon un document publié par le Bureau du vérificateur général (1998), le vérificateur doit maintenant :

- attacher une importance encore plus grande à la vérification axée sur les résultats (par opposition à la vérification axée sur les systèmes) ;
- adopter une démarche fondée sur « les meilleures pratiques » dans ses vérifications en signalant ce qui fonctionne bien de même que ce qui ne fonctionne pas bien ;
- améliorer les notions et les pratiques relatives à la reddition de comptes (notamment la gestion axée sur les résultats, la mesure de rendement et la communication de l'information).

Enfin, plus récemment encore, les ministères fédéraux et le secrétariat du Conseil du Trésor ont commencé, sur une base exploratoire, à élaborer une nouvelle génération de cadres d'évaluation davantage adaptés aux besoins d'information sur le rendement des divers décideurs, y compris les gestionnaires de programme et les chargés de l'élaboration des politiques (Laurendeau, 1999, p. 3). De la même façon, la nouvelle réalité du partenariat en réseau, où rappelons-le différents paliers de gouvernement, les citoyens, le secteur privé et les organismes sans but lucratif sont appelés à collaborer tant lors de la formulation des politiques et des programmes publics que lors de leur mise en œuvre, commence à avoir des répercussions sur la manière dont il faut mener l'évaluation (Delacourt, Lenihan, 1999, p. 113-115).

En somme, durant la dernière décennie plusieurs facteurs sont venus modifier radicalement le contexte dans lequel évoluent les gouvernements canadiens : les nouveaux modes de communication et les nouvelles technologies de l'information intégrant les gouvernements et les entreprises partout dans le monde, objectif de réduction des déficits, une prestation de services publics plus holistique, centrée davantage sur le citoyen, l'évolution de la culture politique qui fait une plus grande place à la population dans la gouvernance (Delacourt et Lenihan, 1999, p. 112).

Malgré tous ces changements, l'évaluation demeure largement sous le monopole de l'exécutif au sein du gouvernement fédéral. Les récents développements dont il a été fait mention ne viennent pas contredire cette position, au contraire. Ainsi, dans ses efforts des dernières années en vue de responsabiliser et d'intégrer davantage les gestionnaires de programmes et les autres parties prenantes aux évaluations et en insistant sur l'utilité qu'elles doivent avoir pour ces gestionnaires, le gouvernement donne une fois de plus le signal que l'évaluation managériale est nécessaire. En refusant, décennie après

décennie, de doter la branche législative de l'État de mécanismes d'évaluation, le gouvernement canadien fait en sorte que l'évaluation managériale est également suffisante, bien que le discours qu'il tient demande à ce que soit abordée l'évaluation de légitimation. En somme, «Les Canadiens, en inscrivant depuis plus de 25 ans maintenant l'évaluation dans les processus de gestion sectoriels, limitent d'entrée de jeu la portée de l'évaluation technique qu'ils pratiquent» (Bégin et coll. 1999, p. 305). Nous reviendrons sur ce dernier point en conclusion. Il convient maintenant d'aborder sommairement les autres lieux de l'évaluation au Canada.

8 L'évaluation hors des ministères fédéraux et du Conseil du Trésor du Canada

L'évaluation au Canada ne se résume évidemment pas à l'évolution de cette fonction telle qu'elle est présentée dans les pages précédentes. D'une part, le gouvernement fédéral réalise quelquefois des évaluations par le biais de commissions royales d'enquête, comités de réflexion, fora divers, etc. D'autre part, les gouvernements des provinces contribuent, eux aussi, à son développement au niveau de leurs propres politiques et programmes qu'ils ont formulés et mis en œuvre durant cette période. Il faut souligner que généralement, comme au palier fédéral, cette fonction y est demeurée largement décentralisée au niveau des ministères et que les vérificateurs généraux des différentes provinces font état de l'importance des évaluations managériales au détriment des évaluations de légitimation dans les dernières décennies.

Il faut également indiquer que plusieurs chercheurs universitaires et des firmes du secteur privé ont bénéficié de contrats d'évaluation des différents paliers de gouvernement. Dans la mesure où ils répondent à des appels d'offres précis, on peut émettre l'hypothèse que l'étendue du questionnement évaluatif des rapports qu'ils produisent se limite alors aux aspects managériaux. Soulignons ici que quelques rares universités canadiennes offrent à leurs étudiants une formation en évaluation de programme.

Par ailleurs, le Canada a été l'un des premiers pays où est apparue une société sans but lucratif ayant comme objectif premier la promotion de l'évaluation de programme. La Société canadienne d'évaluation (SCE) est une association regroupant des membres provenant notamment des gouvernements fédéral, provinciaux et municipaux, des organismes de service, des institutions d'enseignement et des entreprises du secteur privé. Elle regroupe 14 sections

régionales (incluant les 10 provinces et une section de la capitale nationale). Elle a principalement pour objectifs d'offrir une table ronde permettant de discuter de la théorie et de la pratique de l'évaluation au Canada et de promouvoir l'évaluation pour les programmes des secteurs public et privé. La section de la capitale nationale, Ottawa, regroupe majoritairement des fonctionnaires fédéraux, les sections provinciales comptant souvent une forte proportion de fonctionnaires provinciaux. Cela démontre bien les liens étroits entre la SCE et les administrations publiques canadiennes. La SCE et ses différentes sections proposent à leurs membres des activités de formation, des colloques et une conférence annuelle à caractère international. De plus, elle est éditeur de la « Revue canadienne d'évaluation de programme ». La SCE est également préoccupée par les questions de formation des évaluateurs et de rayonnement international de l'évaluation. Ainsi, la professionnalisation de l'évaluation de programme par le biais d'une certification fait actuellement l'objet d'un débat au sein de la Société (Tierhoff, 1999). De même, de récentes publications de la SCE font état de propositions d'arrangements, d'institutions, de partenariats afin de développer et de promouvoir ailleurs dans le monde l'évaluation.

Enfin, il faut indiquer ici l'importance que prend la vérification intégrée, tant dans le secteur privé que dans le secteur public. La Fondation canadienne pour la vérification intégrée (FCVI), qu'appuie sans réserve le Vérificateur général du Canada, joue un rôle de plus en plus important en ce qui concerne les évaluations de type managérial et semble vouloir pousser son expertise du côté de l'évaluation de légitimation.

En somme, les gouvernements des provinces de même que les milieux universitaires et le secteur privé s'intéressent, tout comme le gouvernement fédéral, à l'évaluation depuis plus de 20 ans. Toutefois, comme nous le verrons ci-après, cette profusion des lieux et de l'expertise ne doit pas en faire oublier les limites.

9 Conclusion

Aujourd'hui, dans le domaine de l'évaluation des politiques et programmes publics au Canada, un enjeu majeur apparaît : l'utilité des évaluations. Mais l'utilité pour qui ? On doit se demander si, dans le cadre des arrangements institutionnels décentralisés de l'évaluation qui consacrent à toutes fins pratiques le monopole de l'exécutif sur la fonction, une évaluation peut être

produite et répondre à la fois aux attentes des gestionnaires de programmes, des parlementaires et des citoyens canadiens.

9.1 Les gestionnaires de programme

Les textes récents du secrétariat du Conseil du Trésor que nous avons recensés précédemment indiquent clairement que les gestionnaires de programme doivent être davantage parties prenantes à la formulation même des évaluations à produire. Dans le contexte de partenariat où baigne dorénavant l'exercice d'évaluation auquel ces gestionnaires sont associés: « Les partenaires doivent s'entendre sur les mécanismes à utiliser, comme la vérification interne et l'évaluation de programme pour surveiller les progrès et évaluer le rendement et les résultats. (...) Lorsqu'on établit des mécanismes de collaboration, il est nécessaire d'établir un équilibre entre trois facteurs qui tendent à orienter les efforts dans des directions différentes: la reddition de comptes, l'efficacité et l'atteinte des résultats.» (Desautels, dans Delacourt, et coll., 1999, p. 36-37)

Cet équilibre est difficile à respecter. Le Vérificateur général souligne que l'accent ne devrait pas être uniquement mis sur une plus grande efficacité ou sur la conformité aux exigences de la reddition de comptes comme le font trop souvent les évaluations, sous les pressions des gestionnaires de programme eux-mêmes.

Par ailleurs, un défi particulier concerne les gestionnaires de programmes dont la responsabilité est partagée entre plusieurs ministères. Des propositions sont actuellement à l'étude au secrétariat du Conseil du Trésor pour clarifier le rôle des gestionnaires dans la réalisation d'évaluation dans ce contexte (Laurendeau, 1999). Il est à noter que cette responsabilité de l'évaluation interministérielle incombe au secrétariat du Conseil du Trésor depuis 1977 et que manifestement il n'a pas su encore s'y adresser dans le cadre des arrangements institutionnels en place.

9.2 Les parlementaires

Il semble que, pour l'instant, les parlementaires ne trouvent pas leur compte dans les mécanismes actuels de l'évaluation. Ainsi, lors de consultations récentes, ils ont exprimé l'idée d'augmenter le partenariat entre eux et les bureaucrates, par exemple en confiant un rôle plus important aux comités

parlementaires pour le suivi des rapports de performance (Delacourt et Lenihan, dans Delacourt, 1999, p. 118). Ils ont également indiqué leur intérêt de voir les comités parlementaires "jouer un rôle plus important dans le processus de suivi des évaluations (Thomas, 1999).

Or, une difficulté importante réside dans le fait que la culture des fonctionnaires fait en sorte qu'ils n'ont pas droit à l'erreur, que cela est mal vu. Il faut également rappeler ici la responsabilité ministérielle dont les origines remontent à la Constitution canadienne. Dans cette optique, on comprend que la culture politique ne fait rien pour calmer le jeu. En effet, la nature partisane des débats parlementaires, notamment lors de la période des questions à la Chambre des Communes du Canada ou à l'Assemblée nationale du Québec ramène souvent le portrait qui sera diffusé dans les médias à « tout est noir ou blanc plutôt qu'à accepter des nuances » (Delacourt, Lenihan, dans Delacourt, 1999, p. 113-115). L'on voit mal comment, dans un tel contexte, les fonctionnaires accepteraient de collaborer davantage avec les parlementaires au sein des comités parlementaires ou ailleurs.

9.3 Le citoyen canadien

Il a été démontré que les arrangements institutionnels de l'évaluation de politique et programme au Canada favorisent depuis 25 ans maintenant la multiplication des évaluations managériales au détriment des évaluations de légitimation. Le citoyen pourtant s'intéresse davantage à la question de la raison d'être, des impacts et du rendement économique des politiques et des programmes qu'il défraie à même ses impôts et taxes. Le citoyen veut connaître, peut-être encore davantage aujourd'hui, la réponse à la question de savoir si son gouvernement réalise les bonnes choses, et non pas seulement s'il fait bien les choses qu'il entreprend.

Or, nous avons vu qu'au Canada les évaluations de politique et programme sont souvent réalisées en l'absence de concurrence, ce qui en limite trop souvent la

(1) En premier lieu, ils définissent et communiquent les résultats qu'ils visent ; ensuite ils mesurent leur rendement en améliorant les résultats obtenus au moyen d'une gestion fondée sur les résultats et d'un respect des valeurs primordiales et, finalement ils améliorent les rapports qu'ils présentent au Parlement.

(2) Il est entendu par « dépenses de programme » toutes les dépenses fédérales à l'exception du service de la dette.

portée aux questions d'ordre managérial. Ce manque de concurrence peut s'expliquer par l'orientation administrative de l'évaluation fédérale, c'est-à-dire que l'évaluation a été, est et semble devoir demeurer une fonction décentralisée au niveau des ministères et des organismes fédéraux. Cette réalité fait en sorte que l'évaluation est encore aujourd'hui un outil tourné essentiellement vers les préoccupations de la branche exécutive de l'État canadien, sans que la branche législative ait les leviers nécessaires pour balancer les pouvoirs. Cet état de fait dessert de notre point de vue le citoyen.

Dans le cadre actuel des arrangements institutionnels où l'évaluation demeure une fonction décentralisée au sein des processus de gestion sectoriels de l'État, est-il réaliste de croire que l'équilibre recherché par le Vérificateur général entre la reddition de comptes, l'efficacité et l'atteinte des résultats est possible ? Les Vérificateurs généraux et dans une moindre mesure le secrétariat du Conseil du Trésor du Canada ont relevé, année après année, les limites des évaluations produites dans le cadre monopolistique en place depuis 1977. Dans ce contexte, ne vaudrait-il pas mieux s'engager résolument dans la mise en place d'instances d'évaluation rattachées à la branche législative de l'État ? Le cul-de-sac actuel invite plus que jamais à examiner sérieusement cette option, si réellement on désire que le citoyen soit mieux informé de la pertinence, des effets et du rendement relatif des politiques et programmes canadiens.

10 Références bibliographiques

Agence canadienne de développement international, « La gestion axée sur les résultats à l'ACDI : Guide d'introduction aux concepts et aux principes », Ottawa, Direction de la gestion axée sur les résultats, 1999.

Bégin, (C) ., Bergeron (P), Forest (P-G), Lemieux.(V) , « Le système de santé québécois: un modèle en transformation », Presses de l'université de Montréal, Montréal, 1999, 439 p.

Bégin, (C.), Joubert (P) et Turgeon (J), « L'évaluation dans le domaine de la santé : conceptions, courants de pensée et mise en œuvre », dans Bégin, (C.) et coll., 1999., p. 265-281.

Bryan, (S.), « Mieux mesurer le rendement pour offrir de meilleurs services », *Optimum, La revue de gestion du secteur public*, vol. 27, n° 4, 1997, p. 1-6.

Campbell, (D.T.), et Stanley (J.C.), « Experimental and Quasi-Experimental Designs for Research, Rand McNally College Publishing Co , Chicago, 1963.

Conseil du Trésor du Canada (non daté), Bureau du contrôleur général, « Méthodes d'évaluation des programmes : Mesure et attribution des résultats des programmes » (troisième édition), Approvisionnement et Services, Ottawa, Canada.

Conseil du Trésor du Canada (non daté), « Le point sur l'étude de la fonction d'évaluation », Gouvernement du Canada.

Conseil du Trésor du Canada, « Une gestion axée sur les résultats », Rapport annuel au Parlement du Canada, volume 1, Ottawa, 1999.

Conseil du Trésor du Canada, Rapport ministériel sur le rendement pour la période se terminant le 31 mars 1999, ministre des Travaux publics et Services gouvernementaux, Canada, 1999.

Conseil du Trésor du Canada, rapport ministériel sur le rendement pour la période se terminant le 31 mars 1998, ministre des Travaux publics et Services gouvernementaux, Canada, 1998.

Bureau du Vérificateur général du Canada, secrétariat du Conseil du Trésor ,
« La reddition de comptes dans le secteur public : vers une modernisation »,
document de travail, 6 janvier 1998.

Conseil du Trésor du Canada, rapport sur le rendement, présentation améliorée
des rapports au Parlement, document pilote, ministre des Travaux publics et
Services gouvernementaux, Canada, octobre 1996.

Secrétariat du Conseil du Trésor, « L'évaluation sous la direction d'un
facilitateur - Guide d'assurance de la qualité », Approvisionnement et Services,
Ottawa, 1995a, Canada, 6 p.

Secrétariat du Conseil du Trésor, « Pour une fonction d'examen plus efficace »,
rapport annuel au Parlement par le président du Conseil du Trésor,
Approvisionnement et Services, Ottawa, Canada, 1995b, 50 p.

Secrétariat du Conseil du Trésor, « Le gouvernement fédéral en tant que
« partenaire » : Les six étapes d'une collaboration réussie », Approvisionnement
et Services, Ottawa, Canada, 1995c, 28 p.

Conseil du Trésor , « Manuel for Review », Internal Audit and Evaluation,
Ottawa, 1994.

Secrétariat du Conseil du Trésor, Bureau du contrôleur général, « Votre guide
pour mesurer la satisfaction des clients », Approvisionnement et Services,
Ottawa, Canada, 1992.

Comité sénatorial permanent des finances nationales, Sénat du Canada, « Le
système d'évaluation des programmes du gouvernement du Canada », Ottawa,
janvier 1991.

Secrétariat du Conseil du Trésor, Bureau du contrôleur général, « Méthodes
d'évaluation des programmes », Approvisionnement et Services, Ottawa,
Canada, 1991.

Secrétariat du Conseil du Trésor, Bureau du contrôleur général, « Guide pour
l'établissement du coût des extrants dans le gouvernement du Canada »,
Approvisionnement et Services, Ottawa, Canada, 1989a.

Secrétariat du Conseil du Trésor, Bureau du contrôleur général, « Normes d'évaluation de programmes dans les ministères et organismes fédéraux », Approvisionnement et Services, Ottawa, Canada, 1989b.

« Lignes directrices sur la qualité », (2^e éd.), Statistique, Ottawa, Canada, 1987.

Groupe de travail ministériel chargé de l'examen des programmes, « Introduction au processus d'examen des programmes », Approvisionnement et Services, Ottawa, Canada, 1986.

« L'échantillonnage d'enquête : un guide non mathématique », Approvisionnement et Services, Ottawa, Statistique, Canada, 1983.

Secrétariat du Conseil du Trésor, Bureau du contrôleur général, « Guide sur la fonction de l'évaluation de programme », Approvisionnement et Services, Ottawa, Canada, 1981a.

Secrétariat du Conseil du Trésor, Bureau du contrôleur général, « L'évaluation de programme : une introduction », Approvisionnement et Services, Ottawa, Canada, 1981b.

Secrétariat du Conseil du Trésor, Bureau du contrôleur général, « Principes pour l'évaluation des programmes par les ministères et organismes fédéraux », Approvisionnement et Services, Ottawa, Canada, 1981c.

Commission royale sur la gestion financière et l'imputabilité, rapport final de la Commission royale sur la gestion financière et l'imputabilité, 1979.

« Répertoire de méthodes d'évaluation des erreurs dans les recensements et les enquêtes », Statistique, Catalogue 13.564F, Ottawa, Canada, 1978.

Conseil du Trésor, Bureau du contrôleur général, « Guide d'analyse avantages-coûts », Approvisionnements et Services, Ottawa, Canada, 1976.

Conseils du Fonds canadien d'adaptation et de développement rural, « Outil et guide d'évaluation des capacités », Ottawa, 1999.

Chen, (H.-T.), et Rossi (P.H.), « The Multi-Goal, Theory Driven Approach to Evaluation : A Model Linking Basic and Applied Social Science », *Social Forces*, 59, septembre 1980, p. 106-122.

Delacourt, Susan, et Donald (G) Lenihan , (éd.), « Collaborative Government : Is There a Canadian Way ? », *New Directions*, n°. 6, Institut d'administration publique du Canada, Toronto, 1999, 121 p.

Desautels, Denis (L.), «La reddition de comptes et les nouveaux modes de prestation de services dans l'administration fédérale : certaines conséquences du partage des responsabilités du gouvernement », dans Delacourt, Susan, et coll., p. 30-37.

Dobell,(R) et Zaussman (D), “An Evaluation System for Government : If Politics is Theatre, Then Evaluation is (Mostly) Art”, *Canadian Public Administration*, 24 (3), 1981, p. 404-427.

Éthier, (G.) (sous la direction de), « L'administration publique : diversité de ses problèmes, complexité de sa gestion », Ste-Foy, Presses de l'université du Québec, 1994.

Guba, (E.), et Lincoln (Y.S.), « Fourth Generation Evaluation », Newbury Park, Sage, 1989.

Inwood, (G.J), « Understanding Canadian Public Administration : An Introduction to Theory and Practice », Prentice Hall, Allyn and Bacon, Ontario, Canada, 1999.

Johnson, (A.W.) « Reflections on Administrative Reform in the Government of Canada », Office of the Auditor General, Ottawa, 1992.

Jorjani, (H.), « L'autoroute de l'information : défis et perspectives de l'évaluation des programmes », *Optimum, La revue de gestion du secteur public*, vol. 27, n° 1, 1997, p. 7-13.

Kernaghan, (K.), et Siegel (D), « Public Administration in Canada », Methuen, Toronto, 1987.

Laurendeau, (M), « Cadre de responsabilisation et de rapport sur le rendement – observations acceptées », document de travail, Secrétariat du Conseil du Trésor, Ottawa, 1999, 9 p.

Laurendeau, (M), « Cadres d'évaluation de rapport sur le rendement : présentation à la SCE », Secrétariat du Conseil du Trésor, Ottawa, mai 1999.

Mayne, (J), « Régler la question de l'attribution par l'analyse de la contribution : utiliser judicieusement les mesures de rendement », document de travail, Bureau du vérificateur général du Canada, 1999.

McQueen, (C.) , « Program Evaluation in the Canadian Federal Government », dans Hudson (J), Mayne (J) et. Thomlison (R) (eds), « Action-Oriented Evaluation in Organizations : Canadian Practices, Wall & Emerson, Toronto, 1992, p. 28-47.

McQueen, (C.), « Linking Program Evaluation to Decision Making », *Optimum, La revue de gestion du secteur public*, vol. 15, n° 4, 1984, p. 30-39.

Müller-Clemm, (W)., et Barnes (M.P.), « A Historical Perspective on Federal Program Evaluation in Canada », *La revue Canadienne d'évaluation des programmes*, vol. 12, n°1, 1997, p. 47-70.

Perret, (B), « Le contexte français de l'évaluation : approche comparative », *La revue canadienne d'évaluation de programme*, vol. 9, n° 2, 1994, p. 93-114.

Rossi et Freeman , « A Systematic Approach », Beverly Hills, Californie, Sage, 1979.

Scriven, (M), « Truth and Objectivity in Evaluation », in Chelimsky, (E.), Shadish, (W.R.), (ed.), « Evaluation for the 21st Century - A Handbook », Thousands Oaks, Sage, 1997, p. 477-500.

Scriven, (M), « The Logic of Evaluation », Irviness, Californie., Edgepress, 1980.

Segsworth, (R.V.), « Auditing and Evaluation in the Government of Canada : Some Reflections », *Canadian Journal of Program Evaluation*, ol. 5, n° 1, avril-mai, 1990a, p. 41-56.

Segsworth, (R.V.), « Policy and Program Evaluation in the Government of Canada » dans Rist (R.C.) (ed) « Program Evaluation and the Management of Government », New Brunswick, Transaction, 1990b, p. 21-36.

Saint-Pierre, (M)., Bégin (C), Joubert (P) et Turgeon, (J), « Les arrangements institutionnels de l'évaluation des politiques et programmes sanitaires », dans Bégin, (C.), et coll., p. 283-311.

Thomas, (J), « Improved Reporting to Parliament », *New Directions*, n°. 4), Institut d'administration publique du Canada, Toronto, 1999, 10 p.

Tierhoff, Karen (A), « La certification des évaluateurs de programmes - Une enquête-pilote », septembre 1999, 7 pages.

Turgeon, (J), « Le programme d'action communautaire pour les enfants (PACE) », dans Kessler, (M.-C)., Lascoume, (P)., Setbon, (M)., Thoenig, (J.-C)., *Évaluation des politiques publiques*, Paris, L'Harmattan, 1998, p. 195- 220.

Turgeon, (J)., « Évolution et leçons de l'évaluation de programme aux gouvernements du Canada et du Québec », dans Éthier (G) ., p. 83-108.

Vérificateur général du Canada, « Manuel de vérification de l'optimisation des ressources », Gouvernement du Canada, Ottawa, 2000.

Vérificateur général du Canada, « L'évaluation des nouveaux modes de prestation de services », document de discussion, Gouvernement du Canada, Ottawa, 1999.

Vérificateur général du Canada, « Repenser le rôle de l'État : Un gouvernement pour tous les canadiens », Gouvernement du Canada, Ottawa, 1997.

Vérificateur général du Canada, « Annual Report of the Auditor General of Canada to the House Commons, Gouvernement du Canada, Ottawa, 1993.

Vérificateur général du Canada, « Report of the Auditor General of Canada to the House of Commons Fiscal Years Ended 31 March 1976 », Gouvernement du Canada, Ottawa, 1976.

Weiss, (C.H.), « Where Politics and Evaluation Research Meet », *Evaluation*, vol. 1, n° 3, 1973, p. 37-45.

Contribution VII

Guide des meilleures pratiques à suivre pour l'évaluation par l'OCDE

L'intérêt accordé aux résultats est un élément central des réformes récemment engagées dans le secteur public des pays de l'OCDE. L'évaluation est importante dans un environnement privilégiant l'obtention de résultats parce qu'elle offre un retour d'informations sur l'efficacité, l'efficacités et les performances des politiques publiques, et peut présenter une importance essentielle pour l'amélioration de ces politiques et l'innovation dans ce domaine. Elle contribue donc fondamentalement à une gestion publique responsable.

L'objet de l'évaluation est d'améliorer la prise de décisions à tous les niveaux. Pourtant, les pays semblent souvent n'en faire qu'un usage limité, notamment pour l'adoption de décisions importantes et en situation de redéploiement budgétaire.

L'objectif des lignes directrices présentées ci-après est d'identifier les principales questions et pratiques que les pays membres de l'OCDE devraient prendre en compte lorsqu'ils s'efforcent d'améliorer l'utilisation des évaluations. Ces lignes portent sur la gestion des activités d'évaluation de l'administration publique et des évaluations individuelles, plutôt que sur la méthodologie. Elles n'ont pas pour rôle de déterminer à quel moment l'évaluation peut contribuer le plus aux procédures d'élaboration des politiques et de gestion des performances. Il s'agit là de décisions qui incombent aux pays membres.

OCDE - PUMA Note de synthèse n°. 5
Service de la gestion publique
Mai 1998

Les meilleures pratiques à suivre pour l'évaluation sont présentées dans trois sections

- **OPTIMISER LES ÉVALUATIONS**

définit les évaluations, leurs objectifs, les principaux participants, et les coûts et les avantages ;

1. Définition et objectifs ;
2. Identifier les principaux participants ;

3. Formuler les avantages et les coûts

- **ORGANISER LE CADRE DE L'ÉVALUATION**

examine les pratiques suivies pour améliorer l'organisation et l'utilisation des évaluations dans le secteur public ;

4. Favoriser la culture de l'évaluation ;
5. Gestion stratégique des activités d'évaluation ;
6. Amélioration de la crédibilité

- **RÉALISER DES EVALUATIONS EFFICACES**

examine les pratiques et priorités dans la gestion des évaluations individuelles ;

7. Assurer des liens avec les procédures d'élaboration des décisions ;
8. Choisir l'évaluateur qui convient ;
9. Associer les parties prenantes et diffuser ouvertement les résultats.

1 Optimiser les évaluations

1.1 Définition et objectifs

Les évaluations sont des analyses des résultats des politiques, des organismes ou des programmes publics, menées en vue de dégager des conclusions fiables et utiles. Elles ont pour rôle d'améliorer l'information et de réduire les incertitudes, il est cependant évident que même des évaluations menées avec rigueur font largement appel aux jugements subjectifs. On peut distinguer les évaluations ex-ante (ou examen des politiques) et les évaluations ex-post. Nombre de pratiques discutées ici se réfèrent également à ces deux types, même si leurs objectifs diffèrent.

Les évaluations ont essentiellement pour **objectifs** d'améliorer la prise de décisions, l'affectation des ressources et de mieux rendre compte. Ces objectifs peuvent être atteints en informant l'opinion, en étayant solidement les procédures d'élaboration des grandes décisions et en encourageant les organismes à être en situation permanente d'apprentissage.

Les évaluations doivent s'inscrire dans le cadre d'un **processus plus vaste de gestion des performances**. Elles peuvent les compléter et les améliorer, mais non les remplacer.

1.2 Identifier les principaux participants

Les **évaluations bien menées supposent la collaboration des principaux participants** (évaluateurs, usagers et parties prenantes), sous la direction d'un commissionnaire.

Les **mandataires** sont les organismes qui sollicitent la réalisation d'évaluations. Ils programment l'évaluation, en suivent les progrès, reçoivent le rapport final et établissent les décisions sur des suites à y donner. Il peut s'agir de ministères, de services de l'administration centrale (comme les ministères des Finances, ou des organismes indépendants d'audit et d'évaluation). Dans certains cas, le mandataire peut également être l'évaluateur.

Les **évaluateurs** sont les organismes ou les particuliers qui collectent et analysent les données et portent des jugements de valeur sur les sujets évalués.

Les **utilisateurs** de l'évaluation peuvent être des décideurs, les services du budget, des vérificateurs de comptes, des responsables de politiques, de programmes ou du personnel, des utilisateurs de services, etc.

Les parties prenantes de l'évaluation sont les particuliers ou les organisations qui s'intéressent à la politique ou au programme évalué ainsi qu'aux résultats de l'évaluation. Les parties prenantes et les utilisateurs sont souvent les mêmes.

1.3 Formuler les avantages et les coûts

Les avantages des évaluations devraient l'emporter sur leurs coûts et leurs limitations. Il est possible d'influer sur ces coûts et sur ces avantages en assurant une gestion rigoureuse des évaluations et en choisissant les évaluateurs et les méthodes d'évaluation les mieux adaptés.

Le principal intérêt des évaluations est qu'elles permettent une **étude approfondie des performances** et une évaluation indépendante de l'efficacité d'autres instruments de gestion des performances. Leurs avantages potentiels sont proportionnels à l'importance des politiques ou des programmes visés.

Par ailleurs, l'expérience montre que les évaluations ont tendance à être **trop coûteuses et trop longues**, par rapport à l'usage qui en est fait et à leurs effets. Il y a aussi un risque de voir les évaluations servir à freiner le processus de prise des décisions et justifier l'inaction

2 Organiser le cadre de l'évaluation

2.1 Favoriser la culture de l'évaluation

L'**adhésion** aux activités d'évaluation est attestée par la volonté des hommes politiques, des responsables des politiques et des instances de l'administration

centrale (comme le ministère des Finances), de tirer effectivement parti des conseils pour l'action dégagée des évaluations.

La demande d'évaluation doit être formulée, précisée et coordonnée par les parties prenantes internes et externes. Les évaluations dénuées de l'adhésion des intéressés risquent de rester sans effets. Les obstacles institutionnels aux évaluations comme les résistances internes peuvent être atténués par des consultations destinées à créer un climat de confiance.

Les pouvoirs publics peuvent favoriser l'instauration d'une **culture de l'évaluation** qui encourage l'innovation et l'adaptation à un environnement en mutation. Le message essentiel à faire passer devrait être que, pour continuer à répondre aux besoins, les organisations doivent être en mesure de tirer parti des informations recueillies sur leurs résultats.

La formation et le dialogue, des évaluateurs compétents, des demandeurs bien informés et des utilisateurs enthousiastes et éclairés contribuent tous à l'instauration d'une culture de l'évaluation.

2.2 Gestion stratégique des activités d'évaluation

Les modalités de réalisation des évaluations devraient correspondre aux besoins et aux priorités dans différents domaines d'action des pouvoirs publics. Il peut être souhaitable d'institutionnaliser des procédures systématiques d'évaluation dans des domaines clés d'action où les coûts de collecte des données sont élevés et les informations limitées. Cependant, l'adoption d'une approche plus souple donnera souvent de meilleurs résultats et empêchera que les évaluations ne se transforment en formalités bureaucratiques. L'évaluation des activités communes à plusieurs organismes devrait faire l'objet d'attentions particulières.

Les organismes à l'échelon central jouent un rôle important dans la gestion du processus d'évaluation ; **les évaluations concrètes peuvent cependant être décentralisées** de manière à être menées par différents acteurs à tous les niveaux d'administration. Le développement de compétences d'évaluation dans différentes administrations assure **de l'éventail nécessaire de méthodes et perspectives d'évaluation** (issues par exemple d'évaluateurs internes et externes), et de la conformité de chaque évaluation à sa problématique spécifique liée aux objectifs, à la cible, à la crédibilité et la finalité de l'évaluation.

L'octroi de **crédits spéciaux** pour le financement des évaluations peut constituer une incitation importante à l'évaluation des politiques publiques ; cependant, ils peuvent aussi contribuer à encourager le recours à des évaluations quand d'autres méthodes de gestion des performances pourraient être plus adaptées.

2.3 Améliorer la crédibilité

Le manque de crédibilité risque de compromettre l'utilité des résultats des évaluations. La compétence et la crédibilité de l'évaluateur, l'existence d'un climat de confiance mutuelle entre l'évaluateur et les personnes concernées par l'évaluation, la consultation et la participation des parties prenantes ainsi que les processus de diffusion des résultats sont autant de **facteurs influant sur la crédibilité**.

Les normes professionnelles et éthiques ainsi que la qualité méthodologique de l'évaluation (qui est déterminée par différents facteurs : critères pertinents, données et éléments d'information adéquats, conclusions fiables et clairement exposées) peuvent aussi influencer sur la crédibilité de l'évaluation. L'assurance de qualité et un dialogue franc et ouvert sont de nature à améliorer la crédibilité en permettant de faire ressortir les lacunes éventuelles des évaluations et d'y remédier

3 Réaliser des évaluations efficaces

3.1 Assurer des liens avec les processus de prise de décision

L'évaluation de l'information peut jouer un rôle important dans la **formulation des politiques**, en permettant d'améliorer la qualité des interventions des pouvoirs publics et, dans la procédure budgétaire, en permettant de mieux étayer les priorités et de réaliser des économies. Les évaluations pertinentes portent sur des questions qui sont importantes pour l'action des pouvoirs publics, les aspects budgétaires, la gestion et autres raisons stratégiques

Les **objectifs** de l'évaluation ont une incidence sur le choix des responsables de l'évaluation, de la méthodologie utilisée et de la finalité. Cette finalité devrait être clairement définie. Les évaluations devraient être adaptées aux caractéristiques de la politique étudiée et les méthodes d'évaluation adaptées aux objectifs de l'évaluation, en tenant compte de différentes contraintes comme celles de coûts et de temps. Le fait de prévoir dès le début la réalisation d'évaluations et de définir clairement leurs finalités permettra d'en améliorer l'utilité et d'en faciliter la planification.

La **planification** permet d'améliorer la gestion et la qualité des évaluations. Les demandeurs des évaluations sont responsables de leur planification, notamment de la définition des objectifs visés, des critères adoptés, de la collecte de données et des méthodes suivies. Le choix du moment opportun pour procéder à l'évaluation est important mais le cycle des prises de décisions

est souvent imprévisible et il arrive souvent que les décisions doivent être prises avant que les évaluations n'aient été menées à leur terme

3.2 Choisir les évaluations qui conviennent

Il est bon qu'une organisation procède à une **auto-évaluation** lorsque les principaux objectifs visés sont d'acquérir des connaissances sur cette organisation et d'en améliorer le fonctionnement. Cependant, il peut arriver que les agents concernés n'aient ni le temps ni les compétences nécessaires, que l'éventail des questions à évaluer soit trop limité et que la crédibilité des résultats risque d'être contestée. Il est bon que l'évaluation soit menée par les **organismes centraux** lorsqu'il s'agit d'améliorer les priorités budgétaires, et lorsqu'il est important que l'évaluateur ait des liens étroits avec les processus de prise de décisions. Il est bon que l'évaluation soit menée par des **évaluateurs externes** (organismes de recherche ou conseillers en gestion par exemple), lorsque l'objet de l'évaluation est d'offrir des perspectives nouvelles sur les politiques publiques ou lorsqu'il faut faire appel à des experts en évaluation. Cependant, les évaluateurs externes risquent de n'avoir qu'une connaissance insuffisante des activités fondamentales de la culture de l'organisation évaluée ou du contexte de la politique considérée et n'offrir de ce fait que des évaluations trop théoriques. L'évaluation par des **évaluateurs indépendants** est appropriée lorsqu'il s'agit d'aider à mieux rendre compte et à améliorer la transparence. Cependant, les responsables des politiques voire l'administration dans son ensemble peuvent être réticents à en accepter les résultats et les recommandations. Les audits de performances sont souvent identiques aux évaluations. Leurs caractéristiques incluent l'indépendance de l'auditeur et l'accent porte davantage sur l'obligation de rendre compte que sur une amélioration.

3.3. Associer les parties prenantes et diffuser ouvertement les résultats Les parties prenantes, y compris le personnel, peuvent être désignées pour participer à des commissions d'évaluation ou associées dans le cadre de groupes de direction ou de groupes consultatifs. Les méthodes d'évaluation participatives peuvent servir à dégager des consensus et à susciter chez les intéressés la volonté de s'impliquer dans le processus de changement. Le dialogue avec les utilisateurs et le personnel permet de mieux comprendre leurs besoins et leurs priorités et de mieux y répondre. Les modalités de participation doivent être définies en tenant compte des coûts, des contraintes de temps et des risques d'appropriation inhérents à de tels processus.

Une présentation franche des résultats permet d'accroître la crédibilité de l'évaluation et incite fortement à y donner suite. La

diffusion d'informations sur les réunions et la diffusion des rapports sont utiles pour présenter les conclusions et engager le dialogue à ce sujet. Les jugements et les recommandations, formulés sur la base de critères clairs, attirent l'attention et encouragent l'adoption de mesures pour y donner suite. Dans les jugements formulés, il faudrait s'abstenir de condamner mais plutôt privilégier la recherche des moyens de remédier aux problèmes constatés

4 Au sujet de cette note de synthèse...

Dans le cadre de ses travaux sur la gestion des performances, le secrétariat du PUMA a étudié les procédures d'évaluation menées dans les pays membres afin de définir les enjeux et les pratiques suivies en vue d'améliorer l'usage des évaluations. Les lignes directrices présentées plus haut sont le fruit des enseignements tirés de l'expérience des pays membres. Elles traitent des différents aspects présentant un intérêt pour les organismes de l'administration centrale chargée de définir les stratégies d'évaluation de l'administration publique et pour les responsables des évaluations individuelles.

Ces lignes directrices ont été révisées et acceptées tant par le réseau Gestion des performances du PUMA que par le Comité de la gestion publique. Il convient de souligner ici qu'il n'existe pas une méthode qui serait la seule valable pour organiser et conduire les évaluations. Le choix des méthodes retenues dépendra de plusieurs facteurs, notamment les objectifs de l'évaluation, son rôle dans le cadre plus large de la gestion des performances ainsi que de considérations institutionnelles et politiques.

Le rapport de synthèse intitulé « **Vers de meilleures pratiques de l'évaluation** » est disponible depuis septembre 1998. Pour établir ce rapport, le secrétariat a bénéficié du concours d'un groupe de référence composé de hauts responsables et d'experts d'Australie, du Canada, des États-Unis, de Suède et de la Commission européenne.

Le guide des meilleures pratiques, le rapport de synthèse, ainsi que d'autres informations sur les travaux menés par le PUMA dans le domaine de la gestion et de l'évaluation des performances sont accessibles sur le site Internet PUMA :

[http:// :www.oecd.org/puma/mgmtrest/pac/index.htm](http://www.oecd.org/puma/mgmtrest/pac/index.htm)

Pour toute information complémentaire concernant ces travaux, prendre contact avec :

Sigurdur Helgason / E-mail : sigurdur.helgason@oecd.org

Télécopie (33-1).45 24 87 96

**Liste des membres du Conseil national de l'évaluation
(décret du 26 janvier 1999)**

Personnalités qualifiées :

M. Cousquer (Yves), président d'Aéroports de Paris

Mme Grémion (Catherine), directrice de recherches au CNRS

Professeur Mény (Yves), directeur du Centre Robert-Schuman, Institut universitaire européen

M. Perret (Bernard), chargé de mission au Conseil général des Ponts et Chaussées

M. Rochet (Jean-Charles), professeur à l'université de Toulouse-I

M. Seibel (Claude), directeur de l'animation de la recherche, des études et des statistiques, DARES

Membre du Conseil d'État :

Mme Liebert-Champagne (Monique), conseillère d'État

Membre de la Cour des comptes :

Mme Lamarque (Danièle), présidente de la Chambre régionale des comptes de Haute-Normandie

Membres du Conseil économique et social :

M. Bailly (Jean-Claude), président-directeur général, RATP

M. Bury (Jean-Claude), secrétaire confédéral, CFDT

M. Salustro (Edouard), président d'honneur du Conseil supérieur de l'ordre des experts-comptables, vice-président de la Commission des comptes des services

Maire :

M. Martin (Christian), conseiller régional, maire de Draguignan

Conseiller général :

M. Leroy (Philippe), président du Conseil général de la Moselle

Conseiller régional :

M. Delebarre (Michel), président du Conseil régional du Nord-Pas-de-Calais

Rapporteur général :

Mme Trosa (Sylvie), administrateur territorial

Assistante :

Mme Perry-Roy Christiane, tél. : 01 45 56 51 27