

L'OBSERVATOIRE  
DE L'ADMINISTRATION  
PUBLIQUE

# TELEscope

Revue d'analyse comparée en administration publique

Vol. 13 n°1 printemps - été 2006

## L'ÉVALUATION DE POLITIQUES ET DE PROGRAMMES PUBLICS

- **Vincent Lemieux**  
Évaluation de programmes  
et analyse des politiques
- **Steve Jacob**  
L'évaluation de programmes  
au Québec: un état des lieux
- **Guylaine Doré,  
Richard Marceau**  
L'évaluation de programmes  
à la fonction publique québécoise:  
un profil de compétences requises
- **André Viel**  
L'utilité de l'évaluation: des conditions  
favorables et des pratiques variées
- **Vibeke Normann Andersen**  
L'évaluation de l'Éducation au Danemark
- **Keith Mackay**  
Procédures et systèmes gouvernementaux  
d'évaluation et de gestion de la perfor-  
mance publique: les leçons australiennes
- **Mita Marra**  
Changer les organisations publiques  
de l'intérieur. Le cas de l'évaluation  
à la Banque mondiale
- **Luc Lefebvre, Ivory Yong**  
L'évaluation des programmes d'aide  
au développement conduits par  
la Commission européenne. Les grandes  
synthèses thématiques et par pays

# Sommaire

TÉLESCOPE • Vol. 13 n°1 printemps - été 2006

- III Éditorial  
**Louis Côté, Richard Marceau**
- 1 Évaluation de programmes et analyse des politiques  
**Vincent Lemieux**
- 9 L'évaluation de programmes au Québec : un état des lieux  
**Steve Jacob**
- 19 L'évaluation de programmes à la fonction publique québécoise :  
un profil de compétences requises  
**Guylaine Doré, Richard Marceau**
- 31 L'utilité de l'évaluation : des conditions favorables et des pratiques variées  
**André Viel**
- 38 L'évaluation de l'Éducation au Danemark  
**Vibeke Normann Andersen**
- 48 Procédures et systèmes gouvernementaux d'évaluation  
et de gestion de la performance publique : les leçons australiennes  
**Keith Mackay**
- 64 Changer les organisations publiques de l'intérieur.  
Le cas de l'évaluation à la Banque mondiale  
**Mita Marra**
- 87 L'évaluation des programmes d'aide au développement conduits  
par la Commission européenne. Les grandes synthèses thématiques et par pays  
**Luc Lefebvre, Ivory Yong**

La publication de ce *Télescope* consacré à l'évaluation des programmes et des politiques est le fruit de la collaboration entre deux instituts de recherche de l'École nationale d'administration publique du Québec : L'Observatoire de l'administration publique et le Centre de recherche et d'expertise en évaluation (CREXE) ont en effet mis en commun leur savoir et leurs ressources pour préparer et réaliser ce projet. Ce numéro doit beaucoup à l'implication et au travail de l'équipe du CREXE, **Sylvain Bernier, Constance Olivia Herrera, Johann Lucas Jacob, Caroline Larochelle, Francine Sylvain**, sous la direction du professeur **Richard Marceau**. **Steve Jacob**, professeur au Département de science politique de l'Université Laval a par ailleurs apporté son concours à ce projet.

# Éditorial

Par **Louis Côté**

Professeur à l'École nationale d'administration publique du Québec

Directeur de L'Observatoire de l'administration publique

[louis.cote@enap.ca](mailto:louis.cote@enap.ca)

et

**Richard Marceau**

Professeur à l'École nationale d'administration publique du Québec

Directeur du Centre de recherche et d'expertise en évaluation (CREXE)

[richard.marceau@enap.ca](mailto:richard.marceau@enap.ca)

Chers Lecteurs,

Pour préparer et réaliser ce numéro de *Télescope* consacré à l'évaluation, L'Observatoire de l'administration publique a fait appel au Centre de recherche et d'expertise en évaluation (CREXE) de l'École nationale d'administration publique (ENAP). Sous la direction de Richard Marceau, le CREXE assume un leadership dans le développement, l'application et la promotion de l'évaluation, tant au niveau national qu'international, ainsi que dans la formation de chercheurs et d'analystes en évaluation.

Dans le cadre de cette collaboration entre deux instituts de recherche de l'ENAP, il convient de souligner la contribution particulière de Caroline Larochelle, professionnelle de recherche au CREXE. Il faut se féliciter par ailleurs que Steve Jacob, professeur au Département de science politique de l'Université Laval ait accepté de participer à ce projet. Son aide aura été précieuse.

Dans un passé récent, l'intérêt de l'évaluation n'a pas toujours été reconnu par les décideurs publics quand ils ne faisaient pas, parfois, montre de méfiance à l'égard des buts et des méthodes de la discipline. Familiers des contrôles de conformité *ex ante* sur les procédures, les gestionnaires des programmes publics comprenaient mal la portée d'une évaluation *ex-post* des effets de leurs actions dès lors que les objectifs initiaux avaient été atteints dans le respect scrupuleux des règles.

Bien que son importance fut rappelée dans les principes de toutes les grandes réformes administratives de ces dernières décennies à travers le monde, l'évaluation des politiques et des programmes en fut souvent la négligée. Les temps changent. Elle est en train, aujourd'hui, de rattraper le temps perdu et d'occuper la place qui lui revient, celle de pilier de la gestion axée sur les résultats.

C'est sans aucun doute en effet avec l'adoption, par les administrations publiques de tous les pays, de la gestion axée sur les résultats que se révèle pleinement la nature structurante de l'évaluation et sa contribution fondamentale à la caractérisation de la pertinence, de l'efficacité et de l'efficience des interventions publiques. Et que s'impose du même coup, de toute évidence, la nécessité de maîtriser les concepts, les outils et la pratique d'une évaluation des politiques et des programmes moderne et de qualité.

Or, on ne peut le nier, si le principe de l'évaluation fait consensus, sa pratique pose de nombreux défis aux administrations publiques et requiert une période d'appropriation. Les rapports du *Government Accountability Office* (GAO) aux États-Unis ou ceux des vérificateurs généraux du Canada et du Québec ne taisent pas ces difficultés. On y dénonce, entre autres, l'amalgame sur la nature des résultats à atteindre et plus précisément la tendance à confondre les « extrants » des programmes, c'est-à-dire les biens et les services publics, et leurs répercussions sociales, c'est-à-dire le véritable objectif poursuivi par la gestion axée sur les résultats.

On déplore par ailleurs que les évaluations ne soient que trop rarement intégrées au moment de l'élaboration des programmes publics et des prévisions budgétaires. On observe également, depuis une dizaine d'années, que les décideurs publics ont de plus en plus souvent recours à l'évaluation pour rendre compte de la mise en œuvre et des conséquences des interventions gouvernementales. Ce succès confirme certes l'importance de l'évaluation, mais force est de constater que la disponibilité de rapports d'évaluation au moment du processus de décision ne garantit pas toujours leur utilisation.

La ratification par le Congrès américain du *Government Performance and Results Act* (GPRA) et la mise en place récente, toujours aux États-Unis, d'un *Program Assessment and Rating Tool* (PART) montrent que l'arrimage entre l'évaluation et la décision est possible. Il nécessite quelques ajustements, mais il est le gage d'une information claire et organisée du citoyen sur l'utilisation des fonds publics.

Au Québec enfin, l'étude sur la mise en œuvre de la *Loi sur l'administration publique* (LAP) enseigne d'une part que la définition de cibles et d'objectifs précis et mesurables reste laborieuse et, d'autre part, que le suivi de la gestion et des effets des programmes, à l'aide d'indicateurs pertinents, mérite d'être consolidé. Pour autant, au terme de l'étude ou en commission parlementaire, tous les rapporteurs ont insisté sur le caractère essentiel de l'évaluation de programmes pour encadrer les changements au sein des administrations publiques. Ils ont notamment démontré la contribution de l'évaluation à l'atteinte des objectifs de la LAP et son appoint à la gestion pour être en situation de contrôle. Ainsi, même si nombre d'améliorations doivent être apportées pour que s'instaure une véritable culture d'évaluation continue au sein des administrations publiques, un mouvement favorable à l'évaluation de politiques et de programmes publics se dessine.

Les témoignages des experts que nous avons invités à participer à ce numéro de *Télescope*, attestent de ces préoccupations. Mais ils font surtout état de situations, de réflexions, d'analyses et d'études de cas qui sont autant de reflets des avancées significatives du domaine de l'évaluation des politiques et des programmes. Fidèle à la tradition de *Télescope* et de L'Observatoire de l'administration publique, les exemples sont puisés dans les expériences québécoises, canadiennes et étrangères.

Le numéro s'ouvre avec un article de **Vincent Lemieux** qui explore des voies de rapprochement possibles entre l'analyse des politiques et l'évaluation de programmes, voies qui feraient appel à la notion de régulation appliquée aux politiques et aux programmes. Dans une première étape, il se réfère à Brewer et deLeon, Meny et Thoenig ainsi qu'à Howlett et Ramesh afin de positionner la place de l'évaluation dans l'analyse des politiques. Il considère ensuite les travaux de Rossi, Freeman et Wright, Palumbo et Monnier pour documenter la place de l'analyse des politiques dans l'évaluation de programmes. Cette trame de fond théorique lui sert à formuler les principaux constats qui se font jour au sujet

des relations entre les deux domaines puis à avancer trois propositions pour les inscrire au cœur d'une analyse. Le régime public québécois d'assurance médicaments lui fournit l'occasion d'illustrer son propos.

Depuis l'adoption de la Loi sur l'administration publique et le dépôt des rapports d'étude de la mise en œuvre de cette loi, l'évaluation de programmes a-t-elle connu au Québec un véritable essor au sein des ministères et organismes? Quel est l'état de sa pratique? **Steve Jacob** relate les résultats d'une enquête qui a tenté de répondre à ces questions. Il dresse notamment un bilan du nombre d'évaluations réalisées au gouvernement du Québec et de leurs spécificités. Dans son article, il présente également les facteurs de succès de l'évaluation, la valorisation politique des résultats de l'évaluation et les principales pistes à suivre pour développer l'évaluation de programmes au gouvernement provincial.

Dans l'article suivant, **Guylaine Doré** et **Richard Marceau** s'attachent à dessiner un profil de compétences requises chez les évaluateurs et les évaluatrices de programmes dans l'exercice de leur profession. À partir d'une recension des écrits et d'une analyse du matériel utilisé pour la formation et la dotation, un profil théorique des compétences exigibles des personnes responsables des évaluations de programmes a été créé puis soumis à cinquante-cinq membres de la Société québécoise d'évaluation de programmes (SQEP) qui l'ont validé. Ce texte décrit les principaux résultats de l'étude : un profil en sept dimensions qui met en lumière les compétences essentielles des spécialistes de l'évaluation de programmes.

Sur cette trame de fond, nous sommes invités à nous interroger sur l'utilisation qui est ou devrait être faite des évaluations de programmes en soutien au processus décisionnel. L'article d'**André Viel** se veut une réflexion à propos de l'utilisation possible des informations contenues dans les rapports d'évaluation pour statuer sur l'avenir et l'amélioration des programmes. L'auteur rappelle les conditions favorisant l'utilisation des résultats présentés dans les évaluations pour livrer par la suite quelques réflexions tirées de ses années d'expérience à titre d'administrateur responsable d'unités d'évaluation dans le secteur public québécois. Dans un autre contexte, **Mita Marra** s'est également intéressée à ce thème. En 2002, elle a conduit une recherche sur l'utilisation des évaluations de programmes en appui au processus décisionnel à la Banque mondiale. Dans sa contribution à notre dossier, elle présente les principaux résultats extraits des 50 entrevues semi-structurées et des quatre études de cas menées à cette occasion.

On retrouve en effet au niveau international l'élan que connaît actuellement l'évaluation de politiques et de programmes au Québec. Au Danemark par exemple, un intérêt croissant est porté à l'évaluation en lien avec le souci grandissant d'améliorer la performance du système éducatif. **Vibeke Normann Andersen** consacre son article à l'analyse des outils de réforme qui ont été mis en œuvre afin de rendre plus formelle et transparente l'évaluation dans le réseau danois de l'Éducation. La démarche danoise inclut la création d'un Institut national d'évaluation, compétent pour effectuer des évaluations dans toutes les disciplines de l'éducation. L'adoption, par le Parlement danois, d'une Loi sur la transparence et l'ouverture en éducation constitue un autre volet important de la réforme, car cette loi exige de tous les établissements d'enseignement qu'ils divulguent désormais sur Internet les résultats de leurs performances scolaires. Enfin, le ministère de l'Éducation du Danemark encourage les différents établissements d'enseignement à procéder à des auto-évaluations en vue de faire émerger une culture de l'évaluation en milieu scolaire.

L'article de **Keith Mackay** offre également un bilan de réformes de l'évaluation, mais conduites cette fois dans le secteur public australien. Ces réformes se sont manifestées par la succession de deux générations de systèmes d'évaluation de la performance et de la gestion gouvernementale. La première, entre 1987 et 1997, s'était notamment traduite à travers la décision du gouvernement australien d'instaurer une stratégie d'évaluation formelle afin de fournir une information fondamentale sur la performance des programmes et leurs résultats en soutien au processus décisionnel. Le système d'évaluation de deuxième génération, en vigueur depuis 1997, a abandonné, entre autres, les exigences d'évaluation formelle. L'auteur compare ces deux générations et en tire des enseignements pour l'élaboration éventuelle par d'autres pays d'un cadre d'évaluation de la performance publique inspiré de l'exemple australien.

Enfin, depuis les années 1990, la pratique par la Commission européenne d'une évaluation de l'aide au développement international, connaît un réel essor. **Luc Lefebvre** et **Yvory Yong** expliquent la rapidité de ce développement par la conjonction de trois facteurs dominants: le cadre institutionnel dans lequel s'exerce l'aide internationale, l'opportunité politique naît de la fin de la guerre froide et la tendance générale vers l'adoption de la gestion axée sur les résultats. Les auteurs montrent que l'évaluation des politiques d'aide est une course d'obstacles et qu'elle pose des problèmes méthodologiques complexes et difficiles à résoudre. Pour les spécialistes qu'ils sont, l'adoption de méthodes comparatives et participatives pour évaluer les politiques d'aide au développement mériterait sans doute d'être davantage exploitée.

Les différentes contributions réunies dans ce numéro de *Télescope* dépeignent un panorama varié et actuel des diverses facettes théoriques et pratiques de l'évolution de l'évaluation. Les lecteurs sont instruits de manière pertinente des apports de l'évaluation à la gestion axée sur les résultats et des défis que les administrations publiques ont à relever pour que la fonction *évaluation* puisse donner sa pleine mesure dans cette perspective.

Bonne lecture!

## Appel à communications

### LE SERVICE PUBLIC À L'ÉPREUVE DE LA MONDIALISATION

Dans les États occidentaux, l'idée de « service public » est évolutive. Elle repose sur des valeurs, des croyances et des convictions majoritairement présentes au sein de toute société démocratique, à l'effet que certaines activités doivent échapper, en fonction de leur nature, des objectifs visés et des intérêts en jeu, à l'application de la logique marchande de recherche du profit. Il y a « service public » si les pouvoirs publics estiment, à un moment donné et dans leur champ de responsabilité, qu'un bien ou un service (existant ou nouveau) jugé essentiel ou nécessaire pour tous ne peut être abandonné, de manière satisfaisante, à l'initiative privée et à elle seule.

La production de ces biens et la prestation de ces services d'intérêt général se font alors sous l'égide de l'Administration publique. Celle-ci doit opérer de façon transparente selon des principes légitimes dont le premier garantit à tous les citoyens qui y ont droit un accès à ces biens et services. Ainsi l'Administration publique

concourt-elle à l'équilibre et à la cohésion économique, sociale et culturelle de la société. Cette mise en œuvre s'accompagne toujours de processus de régulation et de contrôle, voire d'une évaluation périodique des moyens et de l'atteinte des objectifs, qui remontent *in fine* jusqu'aux autorités politiques, redevables devant les citoyens qui les ont élus.

Si le « service public » est le produit d'un mode d'organisation sociétale, du lien et du contrat social, ses implications et les modes d'action que celles-ci commandent peuvent être utilement visitées en prenant en compte différents points de vue : historique, juridique, économique, culturel, managérial, etc. Mené dans une perspective comparative internationale, un tel examen peut s'avérer d'autant plus riche d'enseignements que les phénomènes observés aujourd'hui à travers les pays, peuvent permettre d'esquisser les contours d'un avenir possible (réussite, écueil, échec) de pratiques déterminées par des facteurs comparables.

La revue *Télescope* consacrera son numéro thématique du printemps 2007 à l'évolution du service public dans le contexte de la mondialisation. Nous invitons chercheurs et spécialistes à soumettre leur proposition d'article sous la forme d'un court résumé avant le 15 août 2006. Les auteurs dont les propositions auront été retenues devront ensuite remettre le texte complet de l'article avant le 1<sup>er</sup> février 2007.

Publiée quatre fois par an, *Télescope* est une revue d'analyse comparée portant sur des problématiques spécifiques intéressantes les administrations publiques. En mettant en perspective des expériences et des modèles de gestion observés à travers le monde, *Télescope* fait connaître les plus récentes avancées en matière de gouvernance publique. Pour le format des articles, les auteurs sont invités à consulter l'adresse :

<http://www.observatoire.enap.ca>  
à la rubrique publications.

Pour tous renseignements relatifs à ce numéro sur le service public ou pour soumettre des articles, les auteurs doivent prendre contact avec :

#### Marc Cambon

L'Observatoire de l'administration publique  
École nationale d'administration publique du Québec  
555, boulevard Charest-Est  
Québec (Québec) Canada G1K 9E5

Téléphone : (418) 641-3000 p. 6433  
Télécopieur : (418) 641-3057  
Courriel : marc.cambon@enap.ca





# Évaluation de programmes et analyse des politiques

Par **Vincent Lemieux**, professeur émérite à l'Université Laval, Québec • vlemieux@sympatico.ca

On considère généralement que l'analyse des politiques sert à décrire ou à expliquer comment se réalisent les politiques dans leurs différentes phases alors que l'évaluation de programmes vise plutôt à juger si les programmes sont conformes à des valeurs prescrites ou présumées. Dans cette optique, l'évaluation apparaît comme un processus parmi d'autres dans le déroulement des politiques, mais ce processus a ceci de particulier qu'il n'est pas toujours nécessaire à la réalisation d'une politique, contrairement aux processus de l'émergence, de la formulation et de la mise en œuvre.

Dans un premier temps, nous allons examiner trois ouvrages connus en analyse des politiques pour mesurer la place que leurs auteurs réservent à l'évaluation de programmes, puis nous considérerons trois ouvrages sur l'évaluation de programmes pour voir ce qu'il y est dit de l'analyse des politiques.

À partir des principaux constats que nous tirerons de cet exercice, nous poserons quelques conditions à réunir pour resserrer les liens entre l'analyse des politiques et l'évaluation de programmes. Après quoi, nous suggérerons deux cadres conceptuels bien établis qui semblent aptes au resserrement de ces liens et nous les appliquerons brièvement au programme québécois d'assurance médicaments.

## ■ LA PLACE DE L'ÉVALUATION DANS L'ANALYSE DES POLITIQUES

Pour établir la place de l'évaluation dans les ouvrages généraux d'analyse des politiques, nous avons choisi trois ouvrages d'introduction parus entre 1983 et 1995, écrits par des auteurs de nationalité différente. Il s'agit respectivement de deux Américains, Brewer et deLeon (1983), de deux Français, Meny et Thoenig

(1989) et de deux Canadiens, Howlett et Ramesh (1995). Ces ouvrages sont bien connus et largement utilisés comme manuels d'introduction à l'analyse des politiques. Nous allons les présenter brièvement avant d'examiner la place qu'ils font à l'évaluation par rapport aux autres processus par lesquels se réalisent les politiques publiques.

### Brewer et deLeon

L'ouvrage de ces deux auteurs, intitulé *The Foundations of Policy Analysis*, est un des plus représentatifs de l'approche dite des étapes des politiques publiques. Dans son livre sur les théories des politiques, Sabatier (1999) a d'ailleurs fait appel à deLeon pour présenter cette approche.

Brewer et deLeon distinguent six étapes, nommées en anglais *initiation, estimation, selection, implementation, evaluation* et *termination*.

Les deux auteurs estiment que l'évaluation est une étape indispensable dans la réalisation des politiques, même si ses conséquences ne sont pas toujours importantes. Ils ajoutent que l'évaluation peut aussi être considérée comme traversant les autres étapes, dont celle de l'émergence (*initiation*) ou de la ré-émergence d'une politique.

L'évaluation peut être faite à l'intérieur ou à l'extérieur de l'organisation qui sélectionne ou met en œuvre une politique, et différents acteurs peuvent en être responsables. Si on entend l'évaluation au sens large, on peut estimer que les électeurs évaluent les politiques quand ils votent et que la presse et les autres médias contribuent eux aussi à l'évaluation des politiques. Parmi les acteurs concernés par l'évaluation, il y a ceux qui s'y opposent parce qu'ils en craignent les résultats. Il en est ainsi des responsables et des employés d'un programme qui est évalué suite à des critiques formulées publiquement. Les décideurs ne cherchent pas uniquement à obtenir de

l'information au moyen de l'évaluation des politiques, ils cherchent aussi, grâce à elle, à mieux contrôler leur environnement pertinent.

Brewer et deLeon ont donc une vision de l'évaluation qui les conduit à l'articuler avec d'autres étapes de la réalisation des politiques publiques. Ils la perçoivent par ailleurs comme une occasion de jeux de pouvoir entre les acteurs. On retrouve également la relation avec l'analyse des politiques dans leur discussion sur les critères utilisés dans l'évaluation des politiques: l'efficacité, l'équité, l'efficacité et l'adéquation des services.

### **Meny et Thoenig**

L'ouvrage des deux auteurs français est intitulé *Politiques publiques* (1989). Il est divisé en trois parties de nature différente. La deuxième partie, plus longue que les deux autres, est consacrée à l'analyse des politiques publiques. Dans le déroulement de la politique, quatre étapes sont singularisées: l'émergence, la décision, la mise en œuvre et l'évaluation.

Comme dans l'ouvrage de Brewer et deLeon, l'évaluation vient après la mise en œuvre, mais son articulation à l'étape précédente demeure un peu vague. La faiblesse des liens prévus entre les étapes est d'ailleurs un trait général de l'ouvrage de Meny et Thoenig quand on le compare à celui de Brewer et deLeon.

La transition de la mise en œuvre à l'évaluation tiendrait au souhait de l'analyste de recenser les transformations, ou de noter leur absence, qui sont induites par les politiques publiques. L'évaluation n'est cependant pas le privilège exclusif de l'analyste. Dans un système politique, il existe aussi des organismes et des milieux qui exercent une activité d'évaluation. On fait ici référence à la presse, aux corps de contrôle au sein du secteur public, aux groupes de pression, aux associations de consommateurs, etc.

Enfin, les deux auteurs insistent sur les mécanismes cognitifs à l'œuvre dans l'évaluation. Ils se rattachent ainsi à l'approche cognitive des politiques publiques, initiée par leurs collègues français Jobert et Muller (1987).

### **Howlett et Ramesh**

Dans leur ouvrage intitulé *Studying Policy. Policy Cycles and Policy Subsystems* (1995), Howlett et Ramesh s'intéressent à trois dimensions des politiques publiques: les différentes approches théoriques, les acteurs, les institutions et les instruments et enfin les processus par lesquels se réalisent les politiques. Cette dernière dimension occupe la majeure partie de l'ouvrage.

Les processus sont divisés en quatre étapes: la mise en place de l'agenda, la formulation, la prise de décision et l'évaluation à laquelle est rattaché l'apprentissage. Notons que cette séquence est la même que chez Meny et Thoenig et ressemble beaucoup à celle de Brewer et de deLeon. Comme dans les deux autres ouvrages, l'évaluation vient après la mise en œuvre.

Quatre remarques de Howlett et Ramesh sont pertinentes pour notre propos. La première avance que l'évaluation des politiques est dominée par l'approche rationaliste et qu'elle se rattache au *public choice* ou au néo-institutionnalisme. Une deuxième remarque rappelle que l'évaluation, comme les autres étapes des politiques publiques, est une activité politique. Aussi, est-il naïf de croire que l'évaluation soit toujours exercée de façon à révéler les effets d'une politique. Souvent, elle cherche plutôt à voiler certains faits, à parvenir à des conclusions déjà arrêtées ou encore à abolir un programme. La troisième remarque affirme que l'évaluation n'est pas réservée aux évaluateurs officiels, mais qu'elle est le fait de tous les acteurs qui s'intéressent à une politique. Enfin, les deux auteurs notent que l'évaluation des politiques peut donner lieu à de l'apprentissage, ce qui permet d'apporter des ajustements aux buts et aux techniques d'une politique.

### **■ LA PLACE DE L'ANALYSE DES POLITIQUES DANS L'ÉVALUATION DE PROGRAMMES**

Comme dans la section précédente, nous allons examiner trois ouvrages bien connus dans le domaine de l'évaluation de programmes, pour voir, cette fois, la place qu'y tient l'analyse des politiques. Le premier est celui de Rossi et de ses collaborateurs (1979), qui a fait l'objet de plusieurs rééditions et qui est un des plus diffusés à travers le monde. Le deuxième est l'ouvrage collectif publié sous la direction de Palumbo (1987) qui révèle la diversification, sinon

la crise de l'évaluation de programmes à la fin des années 1980. Enfin, nous retenons à côté de ces deux ouvrages américains celui d'un Français, Éric Monnier (1992), qui se situe dans le prolongement de l'ouvrage de Palumbo par la variété des études auxquelles il fait appel.

Après une brève présentation générale de ces ouvrages, nous soulignerons comment ils rattachent l'évaluation à l'analyse des politiques.

### Rossi, Freeman et Wright

Le sous-titre de l'ouvrage de ces trois auteurs (*Evaluation. A Systematic Approach*) promet une approche systématique de l'évaluation et l'organisation du volume ne dément pas cette annonce. Après un chapitre introductif sur l'utilisation de la recherche évaluative, l'ouvrage porte sur quatre types d'activités comprises dans ce type de recherche : tout d'abord la planification et le développement d'un programme ; ensuite la vérification que la mise en œuvre correspond au programme tel qu'il a été défini ; un troisième type d'activités vise à établir si le programme a été efficace dans la poursuite de ses buts ; Enfin la recherche évaluative s'intéresse à l'efficacité, c'est-à-dire aux relations entre les coûts et les avantages d'un programme. Un dernier chapitre porte sur le contexte de la recherche évaluative et en particulier sur ses aspects politiques.

Pour les auteurs, l'évaluation systématique se fonde sur les méthodes et les techniques de recherche des sciences sociales. En cela, la recherche évaluative se distingue de l'évaluation de sens commun pratiquée par la plupart des acteurs sociaux ou encore des évaluations de programmes superficielles et impressionnistes qui ne reposent pas sur l'utilisation de méthodes éprouvées.

Même si Rossi et ses collaborateurs n'utilisent pas toujours le vocabulaire de l'analyse des politiques, les activités de recherche qu'ils identifient peuvent se traduire dans ce vocabulaire. La planification et le développement d'un programme correspondent à la formulation d'une politique, alors que les vérifications concernant la mise en application du programme, son efficacité et son efficacité, renvoient évidemment à la mise en œuvre des politiques. Le lien entre la mise en œuvre et l'évaluation est une fois de plus affirmé.

Le dernier chapitre de l'ouvrage, portant sur le contexte de l'évaluation de programmes, montre que la recherche évaluative est une activité politique. De multiples acteurs sont impliqués dans la recherche évaluative et des conflits surgissent entre eux qu'il importe de réduire le plus possible. Enfin, après quelques considérations sur la temporalité différente de la recherche évaluative et de l'action politique, les auteurs discutent du rôle de l'évaluation dans la prise de décision, ce qui montre bien le lien entre l'évaluation et la « reformulation » d'un programme.

### Palumbo

Le choix que nous faisons de cet ouvrage collectif, intitulé *The Politics of Program Evaluation*, s'explique par la volonté de rendre compte d'une crise de l'évaluation de programmes qui s'est produite au cours des années 1980. L'ouvrage dirigé par Palumbo a la mérite de regrouper quelques-uns des principaux protagonistes des approches concurrentes qui s'affrontèrent au moment de cette crise. Selon notre façon de procéder, nous allons surtout retenir des auteurs qui ont collaboré à l'ouvrage ce qui touche aux relations entre l'évaluation de programmes et l'analyse des politiques

Le directeur de l'ouvrage, Dennis Palumbo, traite du thème principal, c'est à dire celui du caractère politique de l'évaluation de programmes. Pour lui, les résultats de l'évaluation d'un programme, quand ils sont utilisés, ont un caractère politique en ce sens qu'ils deviennent une composante des décisions politiques concernant ce programme. L'évaluation est aussi politique, car elle prend position sur la performance d'un programme, même si elle le fait de façon neutre et non partisane. Dans le chapitre suivant, Carol Weiss est du même avis et note que les considérations politiques touchent de trois manières l'évaluation des politiques et des programmes. Premièrement, les politiques et les programmes sont créés à la suite de décisions politiques. Deuxièmement, les résultats de l'évaluation alimentent les débats politiques. Troisièmement, l'évaluation comporte des jugements implicites de nature politique sur les enjeux, la légitimité et les voies de réforme des programmes. Cette considération est toutefois moins reconnue.

Ces remarques, ainsi que d'autres du même ordre faites par Brown et Wildavsky, montrent que lorsqu'on considère son caractère politique, l'évaluation ne se rattache pas seulement à l'étape de la mise en œuvre, mais aussi à l'émergence, à la formulation et à l'adoption des politiques.

Les interrogations très répandues dans l'ouvrage sur l'objet de l'évaluation et les différentes façons de l'approcher sont révélatrices des débats auxquels l'évaluation de programmes a donné lieu à la fin des années 1980. À côté d'auteurs comme Palumbo, Weiss ainsi que Brown et Wildavsky, qui adoptent des approches plus classiques bien qu'ouvertes aux courants nouveaux, Patton et surtout Guba et Lincoln (1989) proposent des approches constructivistes et plus participatives. Pour eux, l'évaluation résulte davantage des interactions entre les parties prenantes (*stakeholders*), dont les niveaux de savoir sont différents, que de la recherche évaluative de nature systématique, telle qu'elle est présentée dans l'ouvrage de Rossi, Freeman et Wright.

### Monnier

Nous retenons de cet auteur la deuxième édition revue et augmentée de son livre *Évaluations de l'action des pouvoirs publics*, publiée en 1992. Cette deuxième édition a l'avantage de traiter d'études européennes et en particulier françaises, alors que la première, de l'aveu même de l'auteur, se basait surtout sur des études américaines.

La première partie de l'ouvrage porte sur les origines de l'évaluation, notamment aux États-Unis et en France. La deuxième traite de l'évaluation pluraliste, opposée aux modèles américains et français longtemps dominants. Enfin, la troisième partie recense et présente des cas d'évaluation choisis dans différents pays et menés selon différentes méthodes. Dans la conclusion, l'auteur discute des enjeux du développement futur de l'évaluation et des moyens à prendre pour assurer ce développement.

Monnier se situe dans le prolongement de certaines positions prises par des auteurs qui ont contribué à l'ouvrage collectif de Palumbo. Pour lui aussi la démarche évaluative a un caractère politique. Il reprend sur ce point les positions de Carol Weiss et montre que la structure unitaire ou fédérale d'un

pays de même que la division des pouvoirs entre le législatif et l'exécutif font une différence en ce qui concerne le nombre d'évaluations officielles réalisées.

Plus que chez les auteurs du livre de Palumbo, on trouve chez Monnier la conviction que l'évaluation traverse toutes les étapes de la réalisation des politiques, réalisation qu'il qualifie de « tourbillonnaire ». D'une étape à l'autre, les rétroactions et les anticipations provoquées en bonne partie par l'évaluation sont une constante plutôt qu'une exception. Monnier écrit d'ailleurs à la fin de son livre que l'évaluation n'est pas intellectuellement dissociable de l'analyse des politiques. Se demander ce qu'il faut évaluer, ajoute-t-il, conduit à s'interroger sur ce qu'est une politique publique.

Sans prendre ouvertement parti pour l'approche systématique ou pour l'approche constructiviste de l'évaluation, Monnier donne dans son livre plusieurs exemples des deux approches. Il semble cependant se rallier à la conception pluraliste de l'évaluation qui met en présence les différentes parties prenantes, tout en affirmant que le principal enjeu de l'évaluation est sa crédibilité.

## ■ LES PRINCIPAUX CONSTATS SUR LES RELATIONS ENTRE LES DEUX DOMAINES D'ANALYSE

On peut dégager des six ouvrages examinés quatre constats principaux sur les relations entre l'évaluation de programmes et l'analyse des politiques.

1) Dans la plupart des ouvrages, la distinction est faite entre la recherche évaluative de nature savante, qui utilise les méthodes des sciences sociales, et l'évaluation dite courante, pratiquée plus ou moins consciemment par tous les acteurs d'une politique. Pour reprendre un terme proposé par Vickers (1965, p. 39-40), on peut parler d'appréciation des politiques pour désigner ces deux formes de jugements de valeur qui donnent un sens aux jugements de réalité portés par les acteurs.

2) Dans cette conception, l'appréciation est pratiquée à toutes les étapes d'une politique ou d'un programme, et non pas seulement à l'étape désignée de l'évaluation, qui recouvre, en fait, celle des évaluations officielles. On retrouve ce constat dans la plupart des

ouvrages examinés. L'émergence d'une politique ou d'un programme suppose des appréciations de la part des acteurs concernés par les problèmes publics susceptibles de donner lieu à une intervention gouvernementale. La formulation découle souvent d'un processus d'affrontement entre des appréciations contraires émises pendant la recherche des solutions les plus aptes à régler ces problèmes. Quant à la mise en œuvre, elle consiste essentiellement à apprécier l'opportunité des solutions à apporter aux situations considérées comme présentant des problèmes publics.

**3)** L'évaluation de programmes, comme la réalisation des politiques, peut être considérée comme un processus politique, c'est-à-dire comme un ensemble de relations de pouvoir entre les acteurs concernés par la réalisation d'une politique. Tous les ouvrages que nous avons recensés, en particulier celui de Palumbo, soulignent cette caractéristique. Les jugements de valeur sous-jacents à l'appréciation, prise dans le sens de Vickers, sont des ressources importantes du pouvoir, mais il faut bien comprendre qu'elles ne sont pas les seules. Il y a aussi l'information dont dispose un acteur, son expertise, la position hiérarchique qu'il occupe et les leviers de commande en sa possession ainsi que ses liens avec les autres acteurs.

**4)** Dernier constat fait par nombre d'auteurs que nous avons étudiés : il existe dans l'évaluation savante de programmes une séparation entre les partisans d'une approche systématique et ceux d'une approche constructiviste (Marceau et Dubois, 2005). Cette division existe aussi, mais à un degré moindre, dans l'analyse des politiques. Les approches systématiques se fondent le plus souvent sur les théories de l'action rationnelle, alors que les approches constructivistes se fondent sur des démarches herméneutiques où importent avant tout la signification de l'action et l'interprétation qui en est donnée.

## ■ DEUX CADRES CONCEPTUELS APPROPRIÉS

Il y a donc des traits communs à l'analyse des politiques et à l'évaluation de programmes, mais il n'existe pas à l'heure actuelle de cadre conceptuel unifié qui s'applique aux deux domaines d'analyse. À partir des constats qui viennent d'être formulés, on peut estimer qu'un tel cadre conceptuel, qui convienne aussi bien

à l'explication qu'à l'évaluation des politiques, doit remplir trois conditions.

Premièrement, le cadre conceptuel doit faire une place centrale aux valeurs, définies comme des états désirables servant à apprécier des situations de fait. À titre d'exemple, la valeur d'efficacité consiste dans la relation désirable entre les plus grands avantages et les plus petits coûts possibles. Ou bien encore, autre illustration de notre propos, la valeur d'équité réside dans la relation désirable entre des contributions qui tiennent compte des ressources des acteurs et des rétributions qui tiennent compte des insuffisances qui pèsent sur eux. Ajoutons que les valeurs peuvent être également considérées sous l'angle des intérêts, c'est-à-dire de ce qui est utile aux acteurs, ou sous l'angle des croyances, c'est-à-dire de ce qu'ils croient vrai.

Deuxièmement, le cadre conceptuel doit être tel que la présence des valeurs se retrouve à toutes les étapes d'une politique et pas seulement à celle prévue de l'évaluation. Autrement dit, il doit faire une place à l'appréciation dans chacune des phases qu'on distingue généralement dans les politiques publiques : l'émergence, la formulation, la mise en œuvre et l'évaluation.

Troisièmement, un cadre conceptuel propre à resserrer les liens entre l'évaluation de programmes et l'analyse des politiques doit pouvoir rendre compte des interactions entre les acteurs qui participent au déroulement des politiques, y compris à leur évaluation. Ces jeux de politique renvoient aux relations de pouvoir entre les acteurs qui utilisent des atouts pour contrôler les enjeux pertinents aux politiques publiques (Crozier et Friedberg, 1977).

Dans l'ouvrage dirigé par Sabatier (1999) sur les théories ayant pour objet les politiques publiques, on trouve la présentation de sept approches bien établies en analyse des politiques. Notons qu'aucune de ces approches ne se rattache à une vue constructiviste des politiques publiques et que la plupart ne remplissent pas les trois conditions que nous venons de poser. Seules deux d'entre elles les remplissent en partie. Ce sont l'approche des choix rationnels à travers les institutions (*institutional rational choice*) et l'approche des « coalitions plaidantes » (*advocacy coalitions*).

La première de ces deux approches est présentée par sa principale conceptrice, Elinor Ostrom (1999). Dans un schéma qui résume l'approche (p.42), une place explicite est faite à l'évaluation des résultats de l'action et à celle des formes d'interaction entre les acteurs. Au demeurant, plus loin dans le texte, quelques critères d'évaluation sont définis dont l'efficacité, l'équité, l'imputabilité et l'adaptabilité. On retrouve aussi dans le schéma, par le biais des règles institutionnelles qui s'imposent aux acteurs, une dimension politique qui n'est cependant pas formulée en termes de relations de pouvoir. Enfin, de l'avis même d'Edella Schlager (1999, p. 254) dans son bilan des sept cadres conceptuels présentés dans l'ouvrage de Sabatier, l'approche des choix rationnels à travers les institutions permet de couvrir toutes les phases d'une politique publique, contrairement à la plupart des autres approches présentées dans l'ouvrage.

La deuxième approche qui réunit en partie nos trois conditions est celle des « coalitions plaidantes », due à Paul Sabatier et Hank Jenkins-Smith (1999). Selon ces auteurs, les croyances, et en particulier les croyances fondamentales, seraient le ciment des coalitions. Il s'agit là de valeurs concernant la nature des êtres humains, la liberté, la sécurité, etc. Il existe aussi, à un deuxième niveau, des croyances de nature politique et à un troisième niveau des croyances de nature plus instrumentale. Pour les deux auteurs, ces croyances ne sont cependant pas considérées comme des critères d'évaluation des politiques (Turgeon, Gagnon, Bourgault et Garant, 2005). Les « coalitions plaidantes », généralement au nombre de deux à quatre dans un sous-système donné, et les médiateurs (*brokers*) qui cherchent à les rapprocher, sont les principaux acteurs à l'intérieur de cette approche. Pour cette raison, elle apparaît plus politique que celle des choix rationnels dans le cadre des institutions. En revanche, selon Schlager, l'approche des « coalitions plaidantes » couvre moins bien les différentes phases de réalisation d'une politique.

Pour illustrer comment les deux approches pourraient permettre de resserrer les liens entre l'évaluation de programmes et l'analyse des politiques, nous prendrons l'exemple de la politique québécoise de l'assurance médicaments. Après en avoir fait une brève présentation, nous verrons que la place accordée aux valeurs aide à éclairer le comportement des acteurs, que

l'appréciation des situations selon ces valeurs est présente dans toutes les phases de la politique et pas seulement dans les processus officiels d'évaluation et enfin que ces appréciations sont des ressources importantes du pouvoir des acteurs dans le déroulement de la politique. Cet exercice nous offrira aussi l'occasion de comparer les avantages et les inconvénients de chacune des deux approches pour ce qui est du rapprochement entre l'évaluation de programmes et l'analyse des politiques.

## ■ LE CAS DE L'ASSURANCE MÉDICAMENTS

Au cours des années 1990, des rapports d'experts ont soulevé les problèmes rattachés au régime public d'assurance médicaments. Ce régime coexiste au Québec avec les régimes privés d'assurance gérés par les organisations qui assurent une couverture à leurs membres en ce domaine. En 1997, le régime public a été élargi pour assurer une meilleure couverture aux retraités et aux autres personnes qui ne jouissent pas d'un régime privé, dont celles qui n'ont pas d'emploi. Il était prévu que le nouveau régime public exigerait une contribution de la part des bénéficiaires et qu'il serait évalué officiellement après trois ans d'existence (Reinharz, 1999, Lemieux, 2003).

De nombreux acteurs participent aux différents processus par lesquels se réalise le programme. Sans prétendre à l'exhaustivité, on peut énumérer les compagnies pharmaceutiques, Santé Canada qui est l'instance canadienne d'homologation des médicaments, le Conseil québécois du médicament, le ministère québécois de la Santé et des Services sociaux, la Régie de l'Assurance-maladie du Québec, le ministère québécois du Revenu, les médecins, les hôpitaux, les pharmaciens, les représentants des régimes privés d'assurance ainsi que de nombreux groupes de pression, dont ceux qui défendent les intérêts des personnes âgées ou démunies. Le Vérificateur général du Québec (2004) s'est ajouté récemment à cette liste en prenant position dans un de ses rapports annuels.

Ces acteurs sont reliés les uns aux autres dans le déroulement des processus de réalisation du programme. Ainsi les compagnies pharmaceutiques contrôlent la production et la distribution de médicaments, mais encore faut-il que leurs produits soient homologués par Santé Canada et qu'ils soient inscrits

sur la liste du ministère de la Santé et des Services sociaux du Québec suite à une recommandation en ce sens du Conseil québécois du médicament. Le Ministère n'est toutefois pas tenu de suivre les recommandations du Conseil. Les médecins prescrivent les médicaments, livrés ensuite par les pharmaciens, mais ils doivent à l'occasion négocier avec certains de leurs patients, sans compter les pressions qu'ils subissent de la part des compagnies pharmaceutiques pour qu'ils prescrivent certains médicaments plutôt que d'autres. Quant aux groupes communautaires, ils se manifestent surtout quand il y a une évaluation officielle du programme, ce qui est arrivé à plusieurs reprises depuis 1997.

Les valeurs qui inspirent les acteurs sont des éléments significatifs de leurs différences. À cet égard, les valeurs comme critères d'évaluation semblent plus déterminantes et plus opérationnelles que les valeurs comme croyances. Ainsi, on dira que les administrateurs et les experts valorisent surtout l'efficacité, que les groupes communautaires valorisent surtout l'équité, et que les élus valorisent surtout l'imputabilité. Parce que leur contenu demeure flou, les croyances ont une moindre capacité à départager les acteurs entre eux.

En outre, les critères d'évaluation sont plus aptes que les croyances à être appliqués à toutes les phases d'une politique ou d'un programme. Ainsi, les principaux acteurs de la mise en œuvre, soit la Régie de l'Assurance-maladie du Québec, les médecins et les pharmaciens, apparaissent davantage motivés par des critères d'évaluation que par des croyances. Pour autant, il en est peut-être autrement des bénéficiaires.

À l'inverse, l'approche des « coalitions plaidantes » est plus à même que celle des choix rationnels à travers les institutions de rendre compte du caractère politique du programme d'assurance médicaments dans toutes ses étapes, y compris celle de l'évaluation. On peut en effet considérer que le déroulement du programme se joue d'abord entre des coalitions, qui sont des alliances plus ou moins concertées. Les groupes communautaires constituent une véritable coalition, mais le ministre et son entourage ainsi que les administrations (ministères, Régie de l'Assurance-maladie du Québec, Conseil québécois du médicament) forment une alliance beaucoup plus lâche qui correspond assez peu à la définition d'une coalition.

D'autres acteurs comme Santé Canada, les compagnies pharmaceutiques et les représentants des régimes privés d'assurance sont plus ou moins extérieurs à ces alliances. Pour tous ces acteurs, coalisés ou non, les croyances sont des atouts normatifs de leur action et du pouvoir qu'ils exercent. Selon le schéma de Sabatier et Jenkins-Smith, ils disposeraient aussi d'autres ressources, mais elles ont été jusqu'à présent peu étudiées.

L'approche des choix rationnels à travers les institutions ne permet pas de traiter aussi directement du caractère politique du programme, mais elle a l'avantage de montrer que les institutions contraignent ou habilitent les acteurs dans leur recherche du pouvoir.

## ■ CONCLUSION

Nous avons d'abord examiné dans cet article la place faite à l'évaluation dans trois ouvrages d'introduction à l'analyse des politiques, puis la place faite à l'analyse des politiques dans trois ouvrages bien connus sur l'évaluation de programmes. Ensuite, nous avons voulu montrer qu'un cadre conceptuel centré sur les valeurs des acteurs, sur la présence de ces valeurs dans toutes les phases d'une politique et sur le caractère politique de l'évaluation pourrait permettre de resserrer les liens entre l'évaluation de programmes et l'analyse des politiques. À cette fin, deux cadres ont été retenus, celui des choix rationnels à travers les institutions et celui des « coalitions plaidantes ».

Une brève application de ces deux cadres conceptuels au programme québécois d'assurance médicaments a montré qu'ils ont chacun leurs avantages et leurs limites, mais qu'ils pourraient permettre tous deux de mieux intégrer l'évaluation de programmes à l'analyse des politiques.

Chacune de ces deux voies d'analyse tirerait profit d'un tel rapprochement. L'évaluation de programmes s'inscrirait ainsi dans des cadres conceptuels éprouvés en analyse des politiques et celle-ci serait amenée à faire une place plus grande aux valeurs et à l'appréciation des politiques, une dimension incontournable dans l'explication des politiques publiques à chacune des phases de leur réalisation.

## BIBLIOGRAPHIE

- Brewer, G. D., P. deLeon (1983). *The Foundations of Policy Analysis*, Chicago, The Dorsey Press.
- Crozier, M., E. Friedberg (1977). *L'acteur et le système*, Paris, Seuil.
- Guba, E., Y. Lincoln (1989). *Fourth Generation Evaluation*, Newbury Park, Sage.
- Howlett, M., M. Ramesh (1995). *Studying Public Policy. Policy Cycles and Policy Subsystem*, Toronto, Oxford University Press.
- Jobert, B., P. Muller (1987). *L'État en action*, Paris, Presses Universitaires de France.
- Lemieux, V. (2002). *L'étude des politiques publiques. Les acteurs et leur pouvoir*, Sainte-Foy, Les Presses de l'Université Laval.
- Lemieux, V. (2003). « Les politiques publiques et les alliances d'acteurs » dans V.Lemieux et al. (dir.), *Le système de santé au Québec. Organisations, acteurs et enjeux*, Sainte-Foy, Les Presses de l'Université Laval, p.119-143.
- Marceau, R., N. Dubois (2005). « Un état des lieux théoriques de l'évaluation : une discipline à la remorque d'une révolution scientifique qui n'en finit pas », *La Revue canadienne d'évaluation de programme*, vol. 20, printemps, p. 1-36.
- Meny, Y., J. C. Thoenig (1989). *Politiques publiques*, Paris, Presses Universitaires de France.
- Monnier, E. (1992). *Évaluations de l'action des pouvoirs publics*, Paris, Economica.
- Ostrom, E. (1999). « Institutional Rational Choice: An Assessment of the Institutional Analysis and Development Framework » dans P.A. Sabatier (ed) (1999), p.35-71.
- Palumbo, D. J. (ed.) (1987). *The Politics of Program Evaluation*, Newbury Park, Sage.
- Reinharz, D. (1999). « La place du médicament dans le système de santé du Québec », dans C. Bégin et al. (dir.), *Le système de santé québécois. Un modèle en transformation*, Montréal, Les Presses de l'Université de Montréal, p.149-167.
- Rossi, P. H., H. E. Freeman, S. R. Wright (1979). *Evaluation. A Systematic Approach*, Beverly Hill, Sage.
- Sabatier, P. A. (ed.) (1999). *Theories of the Policy Process*, Boulder, Westview Press.
- Sabatier, P. A., H. C. Jenkins-Smith, (1999). « The Advocacy Coalition Framework: An Assessment » in P. A. Sabatier (ed) (1999), p.117-166.
- Schlager, E. (1999). « A Comparaison of Frameworks, Theories and Models of Policy Process » in P. A. Sabatier (ed) (1999), p.233-260.
- Turgeon, J., F. Gagnon, J. Bourgault, S. Garant (2005). « Politiques publiques et santé : les dilemmes de l'évaluation prospective », *Administration publique du Canada*, vol. 48, n° 3, p. 328-347.
- Vérificateur général du Québec (2004). *Rapport à l'Assemblée nationale pour 2003-2004 Tome 2*, p. 15-53.
- Vickers, G. (1965). *The Art of Judgment*, New York, Basic Books.



# L'évaluation de programmes au Québec : un état des lieux

Par **Steve Jacob**, professeur à l'Université Laval, Département de science politique, Québec • [Steve.Jacob@pol.ulaval.ca](mailto:Steve.Jacob@pol.ulaval.ca)

En présentant le rapport général des activités de la Société québécoise d'évaluation de programmes lors de sa dernière assemblée générale, Richard Marceau, président sortant, précisait que « le passé et le présent de l'évaluation au Québec sont riches. Ils nous situent d'emblée au rang des meilleurs. Plusieurs pays nous envient notre passé et notre présent », tout en rappelant que « les choses évoluent rapidement<sup>1</sup> ». Ce bref descriptif de la situation québécoise est partagé par de nombreux acteurs œuvrant dans le milieu de l'évaluation. Sans en contester les fondements, nous souhaitons dans cet article cerner d'une manière plus détaillée le portrait de la situation. À cette fin, nous présentons les résultats d'une étude que nous avons réalisé en vue d'esquisser un état des lieux de la pratique évaluative au sein de l'administration publique québécoise.

Cet article discute successivement du nombre et des caractéristiques des évaluations réalisées, de leurs facteurs de succès, de la valorisation politique de leurs résultats et, finalement, des pistes envisageables pour développer cet outil au niveau provincial. En guise de conclusion, nous formulons quelques premières questions à débattre qui découlent tant des constats actuels que des réflexions qu'ils suscitent en matière de consolidation de l'évaluation de programmes.

Cette présentation permettra de s'interroger sur la pratique actuelle à un moment où l'évaluation constitue

un thème au cœur de la modernisation de l'État. Le gouvernement libéral québécois met en avant les préceptes de gestion de la performance et estime que « la modernisation de l'État implique une évaluation des résultats obtenus. Des mécanismes seront mis en place à cette fin. La diffusion de ces résultats permettra d'informer les citoyens des améliorations effectivement apportées à la qualité des services publics<sup>2</sup> ».

Le plan de modernisation 2004-2007 intitulé *Moderniser l'État. Pour des services de qualité aux citoyens* précise les modalités de réalisation de cet engagement. Ce document met en avant la notion d'évaluation dans une perspective globale puisqu'elle vise tant les organismes dans une perspective d'allègement et de réorganisation structurelle que les programmes. En ce qui concerne ce dernier point, l'objectif poursuivi par une réévaluation des programmes est de mettre fin aux redondances, à la sédimentation et au cloisonnement en vue de s'assurer qu'ils correspondent aux besoins des citoyens. Cette culture de l'évaluation, guidée essentiellement par une quête d'efficacité, doit, selon les vœux de la Présidente du Conseil du trésor, « devenir partie intégrante des modes de fonctionnement de l'État québécois<sup>3</sup> ». Le document consacre un chapitre à la réévaluation des programmes et précise les enjeux qui doivent être pris en considération dans la réalisation de cette entreprise. En vue de systématiser l'évaluation, « le gouvernement annonce la mise en place d'une gestion beaucoup plus cohérente et ordonnée des différents programmes de subventions existant au sein de l'appareil gouvernemental, avec l'instauration d'une véritable « clause crépusculaire » pour introduire une culture de rééva-

<sup>1</sup> Société québécoise d'évaluation de programme, Rapport du président pour l'année 2004-2005, [en ligne], <http://www.sqep.ca/rapports/rapport%20annuel/Rapport04.pdf>

<sup>2</sup> Gouvernement du Québec (2004), *Briller parmi les meilleurs. La vision et les priorités d'action du gouvernement du Québec*, Québec, p. 193.

<sup>3</sup> Gouvernement du Québec (2004), *Moderniser l'État. Pour des services de qualité aux citoyens*, Conseil du trésor, Québec, 2004, p. XI.

luation continue<sup>4</sup> » qui devrait conduire à un examen triennal des programmes. Ceci permettra de rencontrer le désir d'un des répondants de l'enquête qui estime que « tout nouveau programme devrait tenir compte des éléments visant une évaluation future ».

## ■ MÉTHODOLOGIE DE L'ÉTUDE

Nous avons réalisé une enquête écrite auprès des organismes publics au niveau provincial, afin de dresser un état des lieux de l'évaluation des politiques et des programmes, qui se base sur les avis exprimés par les personnes et les institutions directement impliquées. Telle que nous la définissons dans cette recherche, et par conséquent dans le questionnaire de l'enquête, l'évaluation d'une politique publique ou d'un programme consiste en une appréciation rigoureuse réalisée sur la base de différents critères tels que l'efficacité, l'efficacités, la pertinence, etc. Cette démarche ne concerne donc pas les évaluations – souvent individualisées – du personnel de la fonction publique. Pour les répondants à notre enquête, cette définition de l'évaluation semble largement acceptable et acceptée.

Le questionnaire soumis s'articule en trois parties et contient une douzaine de questions. Le premier thème abordé vise à fournir une image de la pratique de l'évaluation au niveau provincial au cours des cinq dernières années. Le deuxième a pour but de mesurer l'utilité perçue de cette pratique. Le troisième tente de dégager les tendances en ce qui concerne un éventuel renforcement, voire une future institutionnalisation, de l'évaluation de programmes.

Le questionnaire de l'enquête a été envoyé par voie postale, de façon personnalisée, le 31 janvier 2005 à 373 personnes que nous avons sélectionnées en vue d'assurer la meilleure représentation possible de tous les ministères et organismes au niveau provincial. Les personnes interrogées se trouvent au sommet de la hiérarchie des institutions concernées. En effet, il s'agit des présidents, directeurs et directeurs généraux, ministres et sous-ministres ainsi que des responsables des services d'évaluation.

Le taux de réponse est de 15,3%. En comparaison avec des exercices similaires, que nous avons réalisés

en Belgique, ce résultat est extrêmement faible. D'un point de vue statistique, il compromet d'ailleurs la représentativité des enseignements de l'enquête. En réalité, cette situation s'explique par une différence fondamentale entre les pays considérés. En effet, en Belgique, les services dédiés exclusivement à l'évaluation de politiques publiques sont extrêmement rares contrairement au Québec et les responsables administratifs répondent donc personnellement à ce genre de sollicitation de recherche. La situation est toute autre au Québec, où bon nombre de responsables administratifs ont fait suivre le questionnaire aux responsables du service évaluatif de leur institution. C'est ainsi par exemple qu'un questionnaire a été rempli une seule fois au « nom » de quatre responsables administratifs ou que quelques responsables administratifs nous ont contacté pour nous signaler qu'ils ne répondraient pas à notre enquête puisque des personnes compétentes en la matière allaient s'en charger. Cette précision nous permet donc, compte tenu de cette spécificité québécoise, de considérer les réponses qui nous sont parvenues comme représentative des ministères et organismes et plus particulièrement du « corps des évaluateurs ». De ce point de vue, la spécialisation de la fonction évaluative est avérée même si certains se demandent « par quelle organisation pourrait-on reconnaître la fonction, la profession de responsable en évaluation de programmes? ».

À la suite de ces précisions, nous devons garder à l'esprit, lors de la lecture et de l'interprétation des résultats, un biais fort probable de notre enquête. Les personnes qui ont pris la peine de répondre au questionnaire représentent celles qui sont les plus sensibles à la question de l'évaluation et qui la pratiquent le plus au sein du secteur public. Les réponses analysées ici peuvent donc donner une image de l'évaluation sensiblement plus favorable que ne l'est la réalité (y compris donc dans les organismes publics qui n'ont pas répondu au questionnaire).

Outre le taux de réponse (qui se base sur le nombre de questionnaires retournés dans les délais), il est également nécessaire de tenir compte de la « qualité » des réponses, c'est-à-dire du soin avec lequel les acteurs ont répondu aux différentes questions. En comparant les questionnaires obtenus en retour par

<sup>4</sup> Gouvernement du Québec (2004), Op. cit., p. 70.

rapport aux questionnaires véritablement exploitables pour notre recherche empirique, il apparaît qu'ils peuvent faire l'objet d'une analyse même s'il y a parfois des différences sensibles sur certaines questions.

## ■ PRÉSENTATION DES PRINCIPAUX RÉSULTATS

Dans les paragraphes suivants, nous présentons en détail les trois thèmes traités par le questionnaire, à savoir la pratique effective de l'évaluation de programmes au niveau provincial, l'utilité de celle-ci du point de vue des acteurs « de terrain » et, finalement, les perspectives pour le futur de l'évaluation.

### La pratique de l'évaluation

#### • Existence et types de pratiques de l'évaluation.

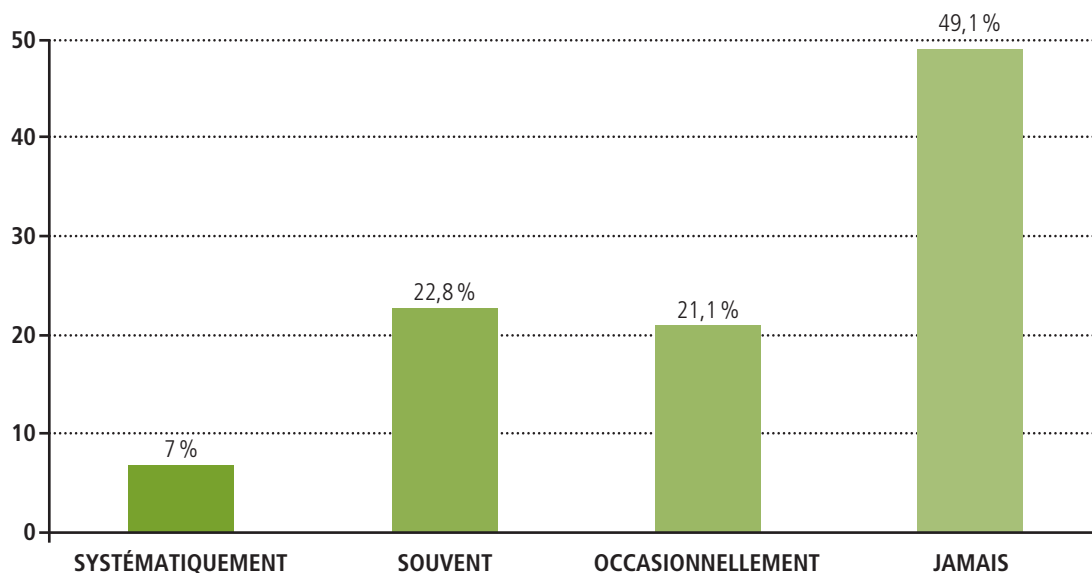
Une courte majorité des répondants (50,9%) déclare avoir réalisé des évaluations au cours des cinq dernières années en comparaison de ceux qui affirment n'en avoir réalisé aucune (49,1%). Ce premier résultat permet de tempérer une impression généralement partagée selon laquelle la pratique évaluative québécoise est pérenne et régulière et que « tout le monde "évalue" »

tout le temps!». Cette situation devrait toutefois évoluer puisque 75% des répondants envisagent de réaliser une évaluation dans le courant de l'année à venir.

L'évaluation n'est pas considérée comme une pratique inutile par ceux qui n'en font jamais et ce n'est pas une méconnaissance des méthodes qui freine le développement de la pratique même s'il arrive parfois que « la méthodologie n'est pas toujours rigoureuse ». Les principales raisons ou justifications invoquées sont essentiellement d'ordre financier<sup>5</sup> (« nous ne disposons pas du budget nécessaire pour financer un évaluateur externe ») et politique (« nous n'avons pas de support/soutien politique »). Comme autres raisons, les répondants précisent qu'ils manquent de ressources humaines dans le service pour conduire des évaluations, que cette pratique ne convient pas pour apprécier toutes les activités publiques ou qu'elle ne figure pas dans le mandat de la direction ou de l'organisme.

La figure 1 ci-dessous, qui présente la répartition de l'ensemble des réponses, démontre que la pratique évaluative est une activité peu systématisée au sein des ministères et organismes québécois.

FIGURE 1 : AU COURS DES CINQ DERNIÈRES ANNÉES, VOTRE SERVICE A PRATIQUÉ L'ÉVALUATION...



<sup>5</sup> Il est vrai que l'évaluation est un exercice coûteux puisque le quart des répondants affirment dépenser plus de 200 000 \$ pour la réalisation d'évaluations.

Le nombre d'évaluations réalisées, sur une période de cinq ans, par une institution, varie considérablement selon les cas (entre 1 et 29 évaluations<sup>6</sup>). Sur la base du nombre d'évaluations accomplies, nous pouvons déduire que ceux qui réalisent « parfois » des évaluations en font, au cours de la période concernée, moins d'une par an. Ceux qui en réalisent « souvent » en font en moyenne trois par an. Ceux qui en réalisent « toujours » en font plus de cinq par an.

Dans la majorité des cas, l'évaluation s'intéresse à un programme d'action (par exemple la mise en valeur des ressources du milieu forestier). Ensuite, l'aspect général semble primer sur le particulier, puisque les évaluations d'une politique publique dans son entièreté (par exemple l'évaluation de la politique environnementale, de santé, etc.) précèdent celles d'une mesure isolée (par exemple l'intervention judiciaire en matière de protection de la jeunesse). Il apparaît donc que l'évaluation vise à expliquer de manière globale un programme d'action ou une politique publique.

Lorsque les services interrogés réalisent des évaluations, les principaux « facteurs de succès » sont, par ordre décroissant, la compétence méthodologique de l'évaluateur ainsi que la participation de l'administration et des usagers au processus évaluatif. Viennent ensuite la clarté des questions définies dans le mandat d'évaluation, l'attitude de la hiérarchie administrative et le choix des techniques utilisées par rapport à la politique évaluée. À l'inverse, les principales difficultés auxquelles se heurtent les ministères et organismes sont le coût trop élevé des évaluations et les délais trop longs avant l'obtention des résultats.

Plus de deux tiers des évaluations sont réalisées *a posteriori*, alors que moins d'un tiers de celles-ci accompagne la mise en œuvre des programmes. Les évaluations *a priori* demeurent vraiment exceptionnelles. Le choix dominant du moment de l'évaluation par rapport au cycle d'une politique peut être lié à la volonté de mieux interpréter et expliquer les effets des programmes d'action, mais aussi à celle de contrôler *a posteriori*, voire de sanctionner, ceux-ci. L'évaluation concomitante permet quant à elle d'amorcer, plus facilement, des processus d'apprentissage collectif et des adaptations incrémentales du programme évalué.

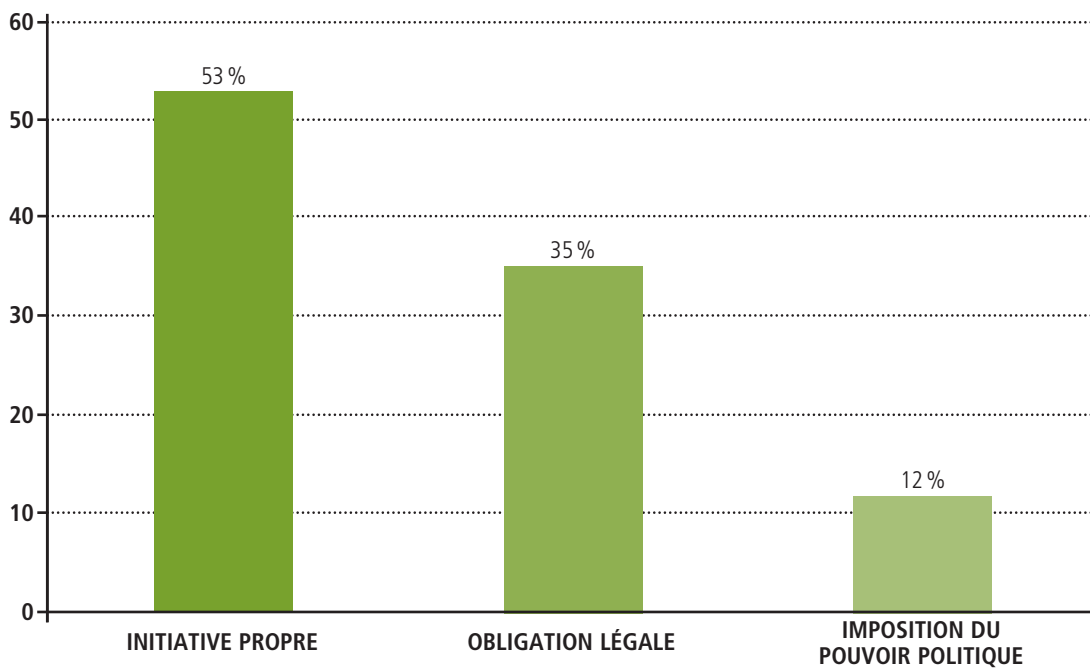
Pour conclure cette section, mentionnons que les répondants s'interrogent sur leurs pratiques évaluatives et que 38 % d'entre eux envisagent de modifier leur approche lors de la réalisation des futurs exercices en vue d'améliorer la planification, notamment en ce qui concerne l'adéquation aux besoins des services concernés par l'évaluation et l'appropriation des résultats produits.

• **Cadre institutionnel de l'évaluation et choix des évaluateurs.** L'évaluation est une pratique à laquelle on recourt pour satisfaire la volonté de certains acteurs administratifs et moins pour se conformer à une obligation légale et/ou réglementaire ou à une injonction politique (cette situation est très rare mais pas inexistante). Ainsi, plus de la moitié des évaluations réalisées résultent d'une initiative interne aux ministères et organismes publics.

---

<sup>6</sup> La phase suivante de notre projet de recherche cherchera à savoir si tous les travaux déclarés par les répondants peuvent effectivement être considérés comme des évaluations (selon la définition présentée plus haut).

FIGURE 2 : LES ÉVALUATIONS SONT INITIÉES PAR...



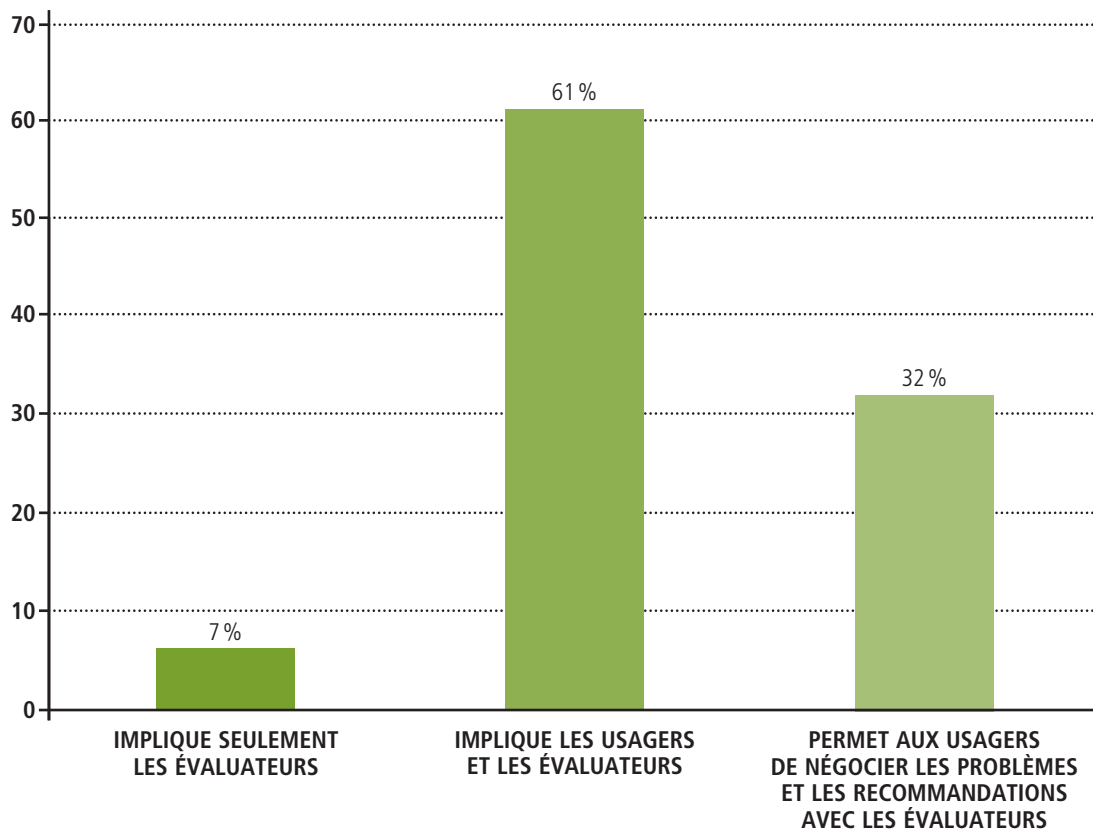
Plus de deux tiers des évaluations sont conduites par un évaluateur appartenant à l'organisation en charge du programme évalué. Lorsqu'ils font appel à un évaluateur externe, les ministères et organismes préfèrent qu'il n'appartienne pas au secteur public. C'est dans ce cas qu'ils s'adressent à un universitaire ou à un consultant privé. Sur la base des réponses obtenues aux questions portant sur l'origine des démarches évaluatives et sur le type d'évaluateur sollicité, nous pouvons faire l'hypothèse que les évaluations lancées à l'initiative de l'organisation sont réalisées par un évaluateur interne, que celles exécutées par obligation légale ou réglementaire imposent l'évaluateur externe et que le pouvoir politique s'adresse à des évaluateurs extérieurs à l'organisation mais issus du secteur public.

Plus de 95% des évaluations se font en intégrant l'administration aux travaux entrepris. Cette proportion s'explique, en partie, par la répartition relative à la

source de la démarche évaluative. En effet, à partir du moment où la majorité des évaluations sont entreprises à l'initiative de l'organisation, il semble logique qu'il y ait une forte interaction avec celle-ci. Par contre, et c'est beaucoup plus surprenant, selon les répondants à cette enquête, dans 80% des cas, les usagers et les parties prenantes sont associés à la démarche évaluative.

L'évaluation québécoise s'affiche donc comme un processus participatif, voire pluraliste. En effet, en plus d'associer des acteurs concernés par le programme évalué, il est intéressant de constater que ces derniers peuvent dans un tiers des cas profiter de cette occasion pour entamer des négociations sur les problèmes rencontrés ou sur les recommandations à formuler.

FIGURE 3 : LE PROCESSUS D'ÉVALUATION...



• **Démarche méthodologique de l'évaluation.** La lecture de manuels théoriques et la participation à des cours de formation continue<sup>7</sup> – notamment à l'ENAP – représentent le mode essentiel d'apprentissage des méthodes et des techniques d'évaluation par les répondants. Il est intéressant de constater que l'expertise en matière d'évaluation des politiques publiques se diffuse également au sein des administrations grâce à un échange d'expériences entre collègues (il semble donc exister une « culture de l'évaluation » partagée au sein du secteur public). Ensuite, il y a une variété de mode d'apprentissage de l'évaluation parmi laquelle figure la participation aux colloques de la Société québécoise d'évaluation de programmes. Malgré cette offre diversifiée, un répondant estime qu'il « faudrait

développer davantage les forums d'échange en matière d'évaluation ».

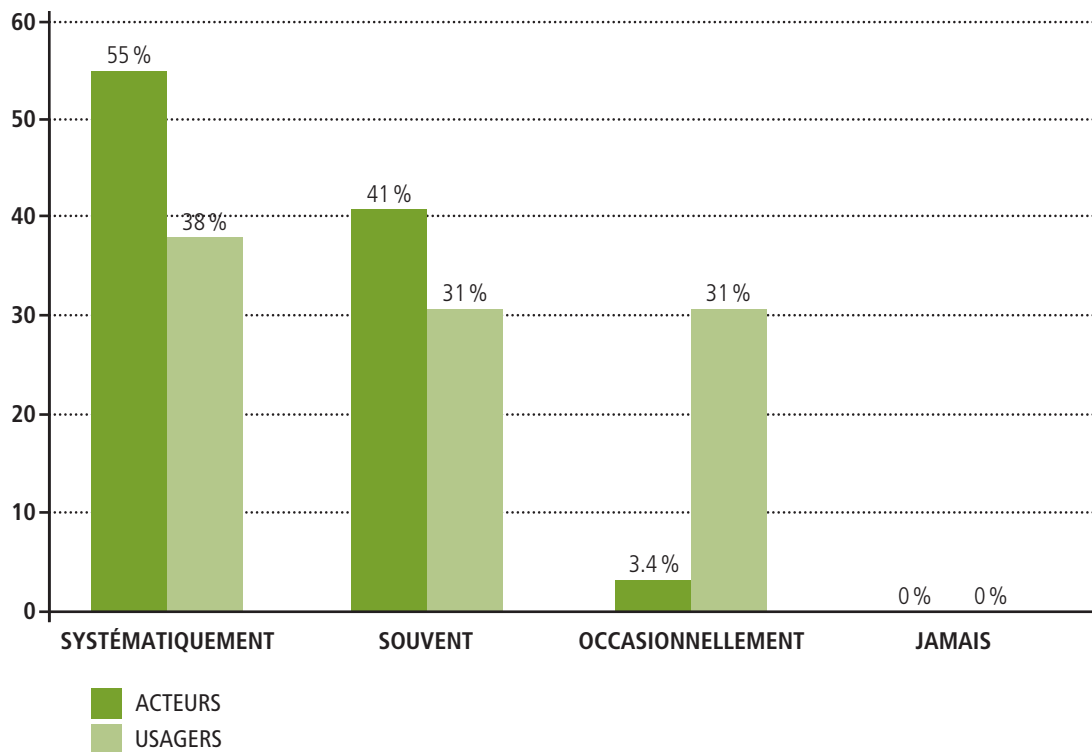
Les méthodes utilisées dans les évaluations de politiques publiques mobilisent des outils et des techniques tant quantitatifs que qualitatifs. D'une manière générale, les sources mobilisées par les services qui réalisent des évaluations mélangent les données primaires (recueillies pour l'occasion) et secondaires (déjà existantes). Les méthodologies les plus fréquemment utilisées lors de la réalisation d'une évaluation sont par ordre décroissant: l'analyse de bases statistiques, les enquêtes par questionnaire, les entretiens individuels approfondis, l'analyse comparative ainsi que les *focus group*.

<sup>7</sup> La formation de base, qui ne figurait pas dans la liste fermée du questionnaire, a également été mentionnée par plusieurs des répondants. Il faut également souligner le fait que plusieurs agents ont rédigé un mémoire et même parfois une thèse de doctorat sur un sujet de nature évaluative.

La dimension comparative est présente dans de nombreuses évaluations réalisées par les ministères et organismes. Le principal mode de contrôle en la matière est soit le niveau fédéral soit les autres provinces et beaucoup moins les comparaisons avant-après la mise en œuvre du programme ou entre les institutions provinciales. De l'avis d'un répondant, des efforts devraient être accomplis afin de « s'ouvrir davantage aux évaluations pluri-ministérielles ».

La figure 4 est particulièrement éclairante et complémentaire de la dimension participative du processus évaluatif que nous avons présenté précédemment. En effet, toutes les évaluations mentionnées permettent aux acteurs (administration et opérateurs) et aux usagers finaux des programmes de participer aux évaluations réalisées.

FIGURE 4 : LES ACTEURS OU USAGERS ONT-ILS L'OCCASION DE PARTICIPER AUX ÉVALUATIONS ?



Toutefois, en croisant cette information avec les résultats sur les méthodes employées, nous nous interrogeons sur la nature de la participation des personnes concernées. En effet, les *focus group* sont marginaux et les entretiens individuels sont loin derrière les enquêtes par questionnaire. Aussi, si la participation se réduit à l'envoi d'un questionnaire, force est de constater que nous n'avons pas véritablement à faire avec un processus participatif au sens classique du terme. Cette remarque vaut plus pour les

usagers finaux que pour les acteurs administratifs. En effet, dès lors qu'une majorité des évaluations sont réalisées sur la base d'une initiative interne et par un évaluateur appartenant au service, il semble logique qu'elle associe les membres des ministères et organismes concernés. C'est également pour cette raison que le rapport est dans près de 80 % des cas transmis au personnel de l'administration.

## L'utilité de l'évaluation pour l'administration

Dans les paragraphes suivants, nous analysons en détail l'utilité des évaluations de programmes telle qu'elle est perçue par les organismes publics ayant participé à cette enquête. D'après les répondants, les évaluations n'ont pas vocation à finir dans un tiroir ou à éclairer uniquement ceux qui les réalisent puisque 29 % des résultats sont systématiquement publiés et que 46 % le sont parfois. Il est rare que le rapport soit considéré comme confidentiel. Environ 20 % des répondants déclarent recourir, systématiquement ou souvent, à cette alternative.

La diffusion du rapport d'évaluation est avant tout organisée en interne, puisque dans 80 % des cas, le rapport est transmis au personnel de l'administration. Les services évalués se prononcent rarement sur le contenu du rapport au moment de sa diffusion. Dans la majorité des cas, il ne contient aucun commentaire de la part des services évalués. Lorsque des observations sont apportées, elles concernent principalement la nature des recommandations formulées ou la qualité générale de l'exercice entrepris.

Il semble donc que les chiffres que nous venons de présenter au sujet de la confidentialité des résultats des exercices évaluatifs sont à relativiser puisqu'ils concernent essentiellement une diffusion en « circuit fermé » à l'intérieur des ministères et organismes. L'évaluation est une pratique qui encourage la « réflexivité » interne de l'administration sans pour autant alimenter un débat public au sujet des actions menées par les organismes publics. Il semble donc que la finalité de gestion prime sur une perspective démocratique. En effet, la transmission aux autres acteurs de la sphère politique est marginale. Seuls 30 % des répondants affirment transmettre systématiquement ou souvent le rapport au ministre, et ils ne sont que 15 % à l'adresser à l'Assemblée nationale. La diffusion sur Internet est systématique (ou presque) dans 50 % des cas. Si cette volonté de transparence est certainement remarquable et à encourager, nous souhaitons tout de même relativiser ces résultats. Chacun sait que face à la masse d'information mise à la disposition des internautes, il est rare qu'une simple mise en ligne suffise à amorcer un débat. Or, puisque les résultats ne font pas souvent l'objet d'une conférence de presse et qu'ils sont rarement soumis à un débat public, l'évaluation québécoise est

essentiellement confinée au sein des administrations et organismes. Tout rapport devrait au moins alimenter la presse spécialisée et éveiller, si ce n'est mobiliser, les évalués. À ce sujet, il est intéressant de constater que la plupart des évaluations sont réalisées sans prendre en considération l'état de l'opinion publique, les préférences des ministres ni les enjeux médiatiques.

Il est bien connu que les retombées d'une évaluation constituent un des enjeux fondamentaux de la démarche. En effet, lors de sa réalisation, une évaluation fait souvent naître sur le terrain des effets de libération, voire des attentes de changement. Il est donc essentiel d'assurer un suivi des conclusions et des recommandations habituellement contenues dans le rapport d'évaluation. Toutefois, ce dernier ne fait pas toujours l'objet de la publicité ou de la diffusion qu'il mériterait, parce que les commanditaires politiques ont été remplacés, parce que les priorités politico-administratives ont changé ou parce que les enseignements de l'évaluation sont trop critiques.

Au niveau provincial, la réappropriation des résultats des évaluations semble assez diffuse selon les répondants. Si aucun répondant ne considère que les enseignements d'une évaluation ne sont jamais pris en compte, force est de constater que les avis sont beaucoup plus partagés entre ceux considérant qu'ils le sont parfois (37 %), souvent (26 %) ou systématiquement (37 %). L'appréciation des effets qui découlent des évaluations est différente selon qu'ils portent sur la politique, sur l'organisation ou sur les relations entre les différents acteurs impliqués dans la conduite de l'action publique. Les répondants estiment majoritairement qu'une évaluation produit des changements importants sur la politique et sur les procédures organisationnelles plutôt que sur les aspects structurels parce que « l'évaluation commence seulement à avoir une influence sur la culture organisationnelle ». Parmi ces changements, figure en bonne place une amélioration de la coordination entre l'administration, les opérateurs et les usagers même si, dans certains cas, il est nécessaire de surmonter l'absence de volonté politique en la matière. À l'inverse, l'évaluation engendre peu de changements sur la connaissance du processus politique si ce ne sont quelques adaptations mineures. À ce sujet, il convient de rappeler que l'appréciation des effets d'une évaluation est subjective et il serait intéressant d'approfondir cette question. En effet, selon un répondant, il est impos-



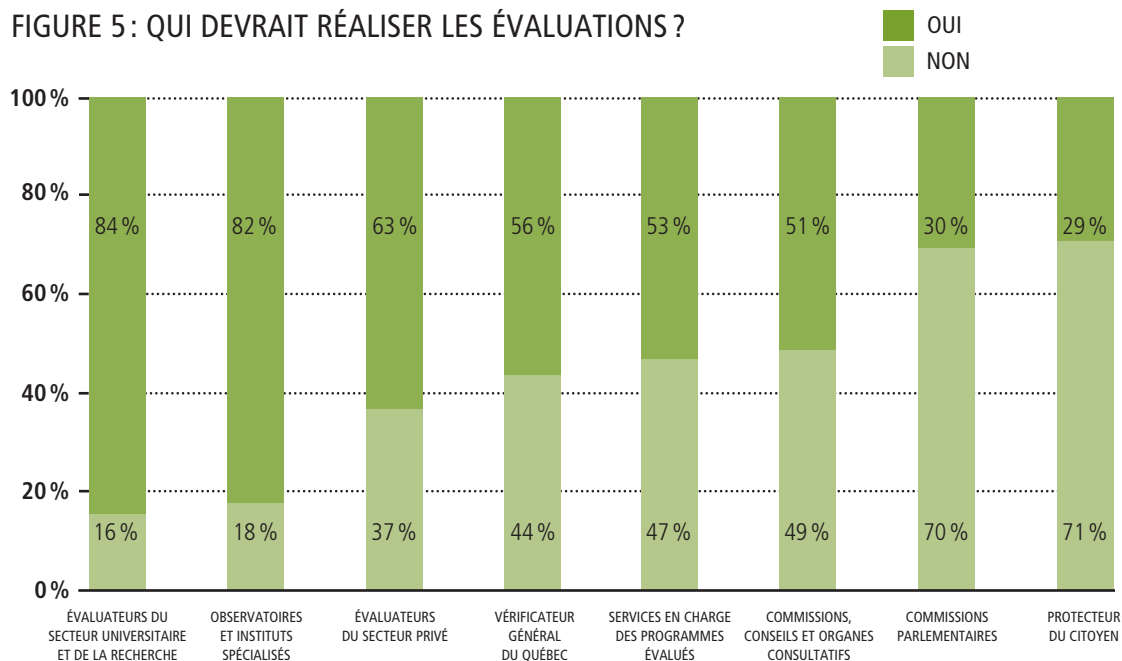
sible de se positionner sur ce point puisque « les rapports ne [lui] appartiennent pas et [qu'il] ne sait pas ce qu'il en advient ».

### Le futur de l'évaluation

Au-delà des pratiques antérieures de l'évaluation et de leur utilité déclarée que nous venons de présenter, nous avons enquêté sur le futur de l'évaluation, en demandant aux acteurs publics de se prononcer sur différentes modalités envisageables pour renforcer institutionnellement l'évaluation au niveau provincial. À ce sujet, nous devons garder à l'esprit le fait que « le développement de l'évaluation est tributaire du type d'État dans lequel il se produit et de la nature des élites susceptibles de la porter, comme des éléments plus culturels à travers lesquels la chose publique trouve sa définition et sa justification<sup>8</sup> ». D'une manière prédictive, mentionnons rapidement qu'une majorité de répondants (68 %) estime que l'évaluation va se développer, au Québec, et que personne ne pense que la pratique évaluative va amorcer un déclin. Indépendamment de cette perception optimiste, il est intéressant de constater que les répondants considèrent à l'unanimité qu'il est nécessaire de développer la pratique évaluative.

Les formes d'institutionnalisation privilégiées par les répondants encouragent un développement externe des capacités évaluatives. Ils estiment qu'il faudrait encourager les évaluations réalisées par le secteur universitaire et de la recherche ainsi que par des observatoires et instituts spécialisés, voire des évaluateurs du secteur privé. Il semble donc se développer une tendance qui consiste à distinguer l'évaluation des activités de mise en œuvre en raison d'une recherche d'un « processus neutre et spécialisé ». Cette volonté d'externalisation de l'évaluation se confirme par le fait qu'une faible majorité de répondants considère que les services en charge des programmes devraient également les évaluer ou qu'il faudrait confier ce mandat au vérificateur général du Québec. Cet élément contredit notre impression présentée ci-dessus au sujet d'une volonté des fonctionnaires d'encadrer le processus évaluatif. Enfin, le rejet d'une évaluation à connotation politique est manifeste puisque seul un tiers des répondants considère que les commissions parlementaires ou le protecteur du citoyen devraient réaliser les évaluations. La figure suivante récapitule ces préférences.

FIGURE 5 : QUI DEVRAIT RÉALISER LES ÉVALUATIONS ?



<sup>8</sup> Duran, Patrice (1993). « Les ambiguïtés politiques de l'évaluation », Pouvoirs, Paris, PUF, n° 67, p. 137.

## ■ EN GUISE DE CONCLUSION

À la lecture des résultats préliminaires de cette enquête, deux éléments méritent une attention particulière pour poursuivre la réflexion sur la pratique évaluative au Québec :

- *Primo*, l'institutionnalisation de la pratique évaluative auprès d'unités distinctes encourage vraisemblablement son développement. Toutefois, il ne semble pas exister de mécanismes qui garantissent ensuite la diffusion des résultats produits dans le processus administratif interne ou auprès des décideurs au sens large. La spécialisation peut dans certain cas engendrer une déconnexion et aboutir à une marginalisation des exercices réalisés.
- *Secundo*, selon les répondants, l'évaluation de programmes alimente le débat public. Toutefois, il ne semble pas exister de stratégie de valorisation des résultats clairement définie. Ainsi, en cédant parfois au culte contemporain de la transparence, il n'est pas certain que la mise en ligne des rapports sur Internet constitue la voie la plus pertinente pour

concourir à une discussion des conclusions et des recommandations. Sur ce point, il serait intéressant de s'interroger sur la communication et la vulgarisation des résultats en vue d'encourager une délibération citoyenne des enseignements. De toute évidence, il est peu vraisemblable de faire l'économie de l'intégration des médias dans une telle stratégie de communication. C'est d'ailleurs un constat partagé par un répondant qui estime qu'il « faut plus [d'évaluations qui soient] plus simples et accessibles aux usagers et plus d'intérêt de la part des médias ».

Ces premiers enseignements font surgir d'autres interrogations qui seront complétées par la suite lorsque, dans le cadre de nos projets de recherche, nous nous attacherons à savoir quelle est la qualité des évaluations réalisées, mesurée à l'aune des standards de la profession ainsi que l'utilité des recommandations formulées.

# L'évaluation de programmes à la fonction publique québécoise : un profil de compétences requises

Par **Guylaine Doré**, Agente de recherche et de planification socioéconomique au ministère de la Santé et des Services sociaux du Québec, Chargée d'enseignement • guylaine.dore@msss.gouv.qc.ca  
et **Richard Marceau**, Professeur titulaire à l'École nationale d'administration publique du Québec • richard.marceau@enap.ca

Depuis plus d'une vingtaine d'années, le gouvernement du Québec emploie des professionnels pour réaliser l'évaluation des programmes et des services des différents ministères et organismes sous sa responsabilité.

Dans le secteur de l'évaluation de programmes, les concours de recrutement ont été conçus et fondés sur la base des exigences spécifiques de chacune des directions concernées. Les critères d'admissibilité exigent généralement de posséder une maîtrise dans le domaine de la mesure et de l'évaluation ou son équivalent et une expérience pertinente dans le domaine de l'évaluation de programmes. Les candidats doivent réussir deux examens de qualification. Le premier, de type analyse-critique, est commun à toutes les personnes qui postulent un poste de professionnel de la fonction publique québécoise. Le second, conditionnel à la réussite du premier, constitue l'épreuve disciplinaire qui porte sur les connaissances théoriques en évaluation de programmes.

Malgré les nombreuses tentatives pour recruter et sélectionner des candidates et des candidats, plusieurs postes demeurent vacants, les listes d'admissibilité pour les postes en évaluation sont vides et le taux de rétention faible.

Nous croyons que l'élaboration d'un profil de compétences permettrait de repérer les compétences attendues, de mieux cerner la profession et d'établir un processus de sélection qui facilitera le recrutement et, conséquemment, favorisera une plus grande rétention des personnes embauchées.

## ■ LA DÉMARCHE ENTREPRISE ET LE CADRE CONCEPTUEL

La démarche, une variante de l'approche « Analyse des tâches », telle qu'élaborée par Fine, dans Berk (1986), comporte trois étapes. En premier lieu, nous avons développé le cadre conceptuel qui dresse l'état des lieux en matière d'évaluation de programmes. La recension des écrits et l'identification des compétences mentionnées dans la littérature et celles fournies par les principaux centres de formation universitaire québécois en évaluation de programmes ont permis de constituer un profil de départ. En second lieu, nous avons transmis l'information colligée à un groupe d'experts, une vingtaine de personnes, pour compléter le profil de compétences. Finalement, le profil de compétences fut soumis à des évaluateurs qui œuvrent dans le domaine afin d'assurer l'exactitude et la pertinence de la forme finale du profil de compétences.

L'usage des profils de compétences pour la dotation des emplois est une pratique relativement récente. Son usage remonte à une vingtaine d'années. Aujourd'hui encore, plusieurs organisations publiques et privées dotent leurs postes sur la base de la description du domaine, des contenus, des mandats et des fonctions confiées. Les concours et les entrevues de sélection permettent de vérifier la concordance entre les attentes formulées dans la description et la formation et l'expérience démontrées par chaque candidat.

Un profil, c'est un agencement ou une articulation de divers éléments ou composantes constituant un ensemble cohérent à l'intérieur d'un métier ou d'une profession. Un profil professionnel décrit les travaux accomplis par tous les travailleurs d'un domaine

professionnel donné (BIT, 1987, repris par Scallon, 2004). Ainsi le profil de compétences décrit l'ensemble des tâches devant être réalisées correctement dans le cadre d'une fonction ou par le titulaire d'un poste précis.

Diverses définitions de compétence sont proposées, de la plus simple à la plus complexe, de la plus opérationnelle à la plus théorique. Viviane De Landsheere (1988) précise simplement qu'il s'agit de « la capacité d'accomplir une tâche d'une façon satisfaisante » tandis que, à l'autre extrême du continuum, Gérard Scallon (2005) signale qu'il existe plusieurs définitions de compétence, selon le contexte, « chacune ayant un caractère définitif » mais qu'on peut toutefois dégager un tronc commun. Parmi les définitions qu'il rapporte, nous avons retenu celles-ci :

**«...[la] capacité qu'a une personne d'accomplir une tâche donnée. Ensemble de connaissances déclaratives, de connaissances procédurales et d'attitudes qui sont activées lors de l'accomplissement d'une tâche donnée.» (Brien, 1997).**

**« La compétence est la possibilité, pour un individu, de mobiliser de manière intériorisée un ensemble intégré de ressources en vue de résoudre une famille de situations-problèmes » (Roegiers, 2000).**

Au sujet de cette dernière définition, Roegiers (2000) précise : « La plupart des définitions de la compétence font mention de problèmes complexes à résoudre. »

Cette notion de problème est cruciale pour bien saisir la notion de compétence. Le problème posé à l'individu doit être relativement nouveau, complexe ou encore comporter des données superflues ou parasites. Perrenoud (1995 et 1999) a clairement indiqué qu'une compétence est en quelque sorte une capacité de mobiliser des ressources cognitives dans une situation de problème.

La conception que Scallon (2000) dégage de ces diverses opinions, plutôt axée sur l'éducation, l'apprentissage et l'approche par compétences, rejoint et complète celles avancées par Guay (1997, 1998), Payette (1992), Marcelin et Bouteiller (1988).

Dans « *Performance et secteurs publics* », Marie-Michèle Guay (1997) rappelle la définition de la compétence donnée par Whelter et Cameron (1991) : « Une habileté

de gestion est un ensemble structuré de comportements séquentiels accomplis dans le but d'atteindre un objectif, de réaliser un résultat désiré ». Dans une étude effectuée à partir de l'analyse des résultats de 400 dossiers de gestionnaires publics, Guay (1997) a utilisé le « *Profil de compétences en gestion* » de l'ENAP du Québec. Elle souligne aussi « l'importance d'identifier les occasions favorisant le développement des compétences relatives à la connaissance de soi, ce que Payette entend par un véritable savoir-faire... » (cité dans Guay, 1997).

Bordeleau (1992, 1993) précise que « les profils de compétences sont formulés en termes d'aptitudes, de capacités, d'attitudes ou de comportements plus ou moins génériques ou spécialisés ». Ce sont des caractéristiques générales requises par l'individu pour être performant. Bouteiller et Marcellin (1998) considèrent qu'il est important de parler des enjeux dominants « pour identifier les compétences clefs ». Ces auteurs dégagent « cinq familles de compétences stratégiques interdépendantes et complémentaires » qu'ils analysent à trois niveaux : 1) la complémentarité entre les compétences professionnelles, 2) la complémentarité entre les compétences individuelles, et 3) le partage des rôles et du pouvoir.

L'usage du profil des compétences est donc une pratique spécifique et réfère à des concepts et à des acquis qui dépassent les composantes de la description d'un poste. Payette (1992) souligne que la préoccupation première, lors de l'élaboration d'un profil de compétences des personnes œuvrant dans un domaine spécifique, consiste à définir un modèle, c'est-à-dire le prototype d'une personne efficace dans ce domaine. Il nous met toutefois en garde contre l'utopie de croire qu'une liste des caractéristiques de cette personne pourrait facilement constituer un portrait distinct et spécifique. Il défend plutôt la thèse de la nécessité de s'attacher à quatre prémisses : la recherche d'un modèle général, l'évaluation concrète d'un poste particulier, la possession de compétences et l'efficacité (Payette, 1992). Comme Payette à propos des gestionnaires, la question que l'on doit se poser est : « qu'est-ce qui distingue le « compétent » de l'incompétent ? » Nous devons ainsi nous interroger sur ce qui caractérise la personne compétente et chercher à déterminer son profil.

## ■ L'ÉLABORATION D'UN PROFIL DES COMPÉTENCES

Tziner et ses collaborateurs (1993) préconisent trois grandes approches qui permettent de dépeindre un profil de compétences : celle dite orientée vers le travailleur (*worker oriented*), celle orientée vers le produit des opérations (*outcome oriented*) et finalement, celle des traits ou habiletés.

Nous avons retenu la vision de Fine (1986) qui emprunte à ces trois grandes approches pour proposer un processus systématique d'élaboration d'un profil de compétences en misant à la fois sur les connaissances théoriques et professionnelles (recension des écrits et des ouvrages pertinents), sur une démarche itérative réalisée par un comité d'experts soigneusement sélectionnés et sur la validation du produit par un large échantillon d'experts.

Le profil de compétences ainsi obtenu permet de couvrir l'ensemble des concepts qui ressortent généralement d'une analyse des tâches de même que les indicateurs et les moyens de mesure implicites. Il se distingue d'autres profils car il est basé sur une analyse des dimensions d'un emploi, ce qui rend invraisemblable l'existence d'une compétence unique ou globale. Le profil des compétences ainsi créé ne peut être décrit par une seule mesure et par un seul score. Il est composé d'une liste de tâches et d'habiletés, de savoirs, de savoir-faire et de savoir-être qui constitue la base du profil spécifique à un type de spécialistes. Autrement dit, il existe une relation directe entre la spécialité et les tâches qui incombent habituellement à la personne type (spécialiste) qui occupe un poste dans le domaine.

Soulignons aussi que le processus proposé par Fine met à profit diverses sources d'information : la documentation scientifique et professionnelle, les comités d'experts, etc. Il utilise des méthodes de collecte et d'analyses qualitative et quantitative.

## ■ LES RÉSULTATS DES TROIS ÉTAPES DE L'ÉLABORATION DU PROFIL DES COMPÉTENCES

Notre démarche comportait trois étapes principales :

1) l'élaboration d'un profil théorique des compétences, 2) la complémentation du profil par un

groupe d'experts dans le domaine et 3) la validation du profil par des spécialistes de l'évaluation de programmes au Québec.

### L'élaboration d'un profil théorique des compétences

Comme les spécialistes en évaluation de programmes de la fonction publique québécoise sont généralement formés au Québec, nous avons examiné la formation qui y est offerte. Deux établissements offrent ces programmes : l'École nationale d'administration publique (ENAP) et l'Université Laval. Les activités et les contenus des formations, annoncés sur les sites WEB des établissements au mois de novembre 2002, ont été examinés. Cinq dimensions spécifiques et une dimension générale peuvent être ainsi relevées : 1) l'évaluation des programmes et l'évaluation institutionnelle, 2) les fondements et les enjeux, 3) les méthodes de recherche et d'analyse des données qualitatives et quantitatives, 4) la formation en science de l'administration avec des contenus en finances, gestion, comptabilité et économie, 5) la mesure et, finalement, 6) les autres sujets connexes.

On retrouve dans ces formations quatre types de connaissances théoriques spécifiques : 1) l'évaluation de programmes, 2) les méthodes de recherche, 3) les statistiques et 4) l'analyse et le traitement des données. L'intégration de ces contenus et la pratique de l'évaluation de programmes sont des conditions essentielles au développement de réelles compétences et s'effectuent dans le cadre d'activités telles que les cours sur la consultation, les études de cas et la rédaction de mémoire, d'essai ou de stage.

• **La Société Canadienne d'Évaluation (SCE).** Les membres de la SCE que nous avons consultés distinguent l'évaluation de programmes d'autres activités professionnelles par l'objectif, l'approche, les méthodes et les « extrants ». Ainsi, l'évaluation aurait pour objectifs :

- d'aborder des questions pratiques et d'aider à la prise de décisions ;
- De porter des jugements sur le mérite ou la valeur d'un programme. À noter ici que certains évaluateurs soutiennent que l'objectif de l'évaluation est de fournir des renseignements nécessaires pour porter de tels jugements, et non de les porter ;

- De chercher des explications, d'essayer de déterminer pourquoi les choses fonctionnent dans différents contextes et pour différentes personnes. L'évaluation essaie d'atteindre la substance d'un programme au lieu de simplement rester à la surface des processus ou de s'en tenir à des mesures superficielles.

En termes d'approches et de méthodes, la SCE énonce que :

- «L'évaluateur applique de façon appropriée des principes de conception de recherche dans des cadres non contrôlés.
- Les données de l'évaluation sont recueillies et traitées de façon systématique.
- En plus d'utiliser des données de programme existantes, l'évaluateur recueille de nouveaux renseignements en vue de répondre à des questions précises au sujet d'un programme.
- L'évaluation doit préférentiellement être détaillée et l'évaluateur examine de nombreux aspects du programme afin de bien comprendre ce qui se passe.
- L'évaluation est pluridisciplinaire et suppose l'utilisation des méthodes provenant de plusieurs domaines d'étude différents.
- Finalement, en termes « d'extrants », une évaluation devrait fournir des données crédibles, fiables, valides et objectives (non biaisées) portant sur la réalisation des objectifs du programme, sur l'attribution des changements apportés (impacts), sur la valeur d'un programme, sur ce qui fonctionne ou non, sur les mesures à prendre pour l'améliorer, sur le pourquoi un programme est mis en œuvre et sur le comment un programme est réalisé ».

La SCE souligne aussi que « Bien que d'autres professions puissent partager certaines caractéristiques, c'est peut-être la combinaison des caractéristiques qui distingue l'évaluation des autres activités ou professions ».

Elle s'est enfin intéressée à établir les connaissances de base requises de l'évaluateur pour remplir avec succès ses activités professionnelles. Elle a déterminé six composants ou éléments – planification et

conception des évaluations, collecte des données, analyse et interprétation des données, aptitudes interpersonnelles et à la communication, gestion de projets et éthique – qui chacun soutient des compétences de différents niveaux. Chacune de ses 23 compétences sont reliées ou au contraire supposent des activités ou des tâches appropriées.

• **La Société Française d'Évaluation (SFE) et l'European Evaluation Society (EES).** Sous la direction de la doctorante Claire Tourmen et du professeur Maurice Basle, un groupe de chercheurs représentant la SFE et l'EES ont développé un référentiel des métiers de l'évaluation des politiques publiques, un champ d'intervention très étroitement associé à celui de l'évaluation de programmes. Le document (mars 2005) est annoncé comme étant en gestation et, conséquemment, il ne reflète pas une position définitive.

Selon ses auteurs, on doit distinguer les tâches réalisées par le responsable de l'évaluation et celles effectuées par celui qui est chargé d'une évaluation. Leur modèle est de type « monotonique », au sens où rien ne peut être interverti sans altérer la logique du processus. Il est aussi hiérarchique de par la condition essentielle de l'accomplissement d'une tâche pour accéder à une étape supérieure.

Le groupe de travail classe neuf compétences selon qu'elles soient cognitives (4), opérationnelles (3) ou sociales (2). Les quatre compétences cognitives sont les connaissances des politiques publiques, de l'expertise sectorielle, du jeu de rôle entre acteurs et des principes méthodologiques. Les trois compétences opérationnelles concernent la connaissance des bases de données, le maniement des outils de recueil et d'analyse des données et l'utilisation des outils de travail collectif. Finalement, les deux compétences sociales se rapportent à la maîtrise de la relation commerciale et à celle de la relation aux élus et aux services opérationnels.

Ce même groupe de travail propose des activités voisines ou complémentaires à celles de l'évaluation. On y retrouve le « *monitoring* », la préparation du budget, l'élaboration, le suivi et la prospective de politiques publiques, le contrôle de gestion, l'assistance à la maîtrise d'ouvrage, l'audit de structure et le conseil en management public.

Il a élaboré un profil général de compétences présentant l'ensemble des principales tâches qui peuvent être confiées à un responsable et à un chargé d'évaluation mais, comme il le souligne, il s'agit de l'amorce d'un travail à bonifier et à valider.

• **La position d'un groupe de chercheurs américains sous la direction de Stevahn.** Stevahn, King et al.

(*American Journal of Evaluation*, 2005) ont présenté la deuxième version d'une taxinomie des compétences essentielles d'un évaluateur de programmes qui s'inspirait des travaux réalisés en 2001 et de ceux du Joint Committee (*Program Evaluation Standards*, 1994), de AEA (*Guiding Principles*, 1995) et de la Société canadienne d'évaluation (*Essential Skills Series*, 1999). Le tableau suivant résume cette taxinomie.

**TABLEAU 1 : LA TAXINOMIE DES COMPÉTENCES ESSENTIELLES DES ÉVALUATEURS DE PROGRAMMES [TRADUCTION LIBRE]**

DIMENSIONS	COMPÉTENCES ESSENTIELLES
1.0 Pratique professionnelle	pour 6 compétences essentielles : 4 capacités et 2 savoirs
2.0 Enquête systématique	pour 20 compétences essentielles : 15 comportements et 5 connaissances
3.0 Analyse de la situation	pour 12 compétences essentielles : 12 comportements
4.0 Gestion de projet	pour 11 compétences essentielles : 11 comportements
5.0 Pratique réfléchie	pour 5 compétences essentielles : 3 comportements et 2 savoirs
6.0 Capacités interpersonnelles	pour 6 compétences essentielles : 4 utilisations, 1 source et 1 démonstration

Tout comme pour le modèle de la Société française d'évaluation (SFE), la taxinomie proposée par King et al. est un modèle multifactoriel et hiérarchique ce qui implique que la maîtrise d'une compétence est une condition à la maîtrise d'une autre.

Cependant, les auteurs précisent que leur taxinomie est en perpétuelle évolution, que ce modèle demande à être validé et que divers niveaux de performances associées aux diverses tâches, encore non élaborées, doivent être établis.

• **Les compétences proposées par la documentation scientifique.** Le domaine de l'évaluation de programmes distingue trois grandes écoles de pensée ou approches : les formalistes ou classiques, les tenants de l'approche fondée sur la théorie des programmes et les naturalistes ou constructivistes.

Les auteurs partisans de chacune des trois approches de l'évaluation proposent des grandes compétences

essentielles pour l'accomplissement heureux ou satisfaisant des tâches de l'évaluateur de programmes. Quoique les trois approches examinées supposent des profils de compétences qui, à priori, sont quelque peu différents, nous estimons que d'une approche à l'autre, les éléments qui composent le profil sont sensiblement les mêmes en termes de contenu. Ce sont des compétences reliées à l'évaluation des besoins, des possibilités d'action, de l'implantation, du processus, des « extrants », des effets, des impacts et du rendement. Cependant, le niveau de maîtrise de chacune varie parfois significativement selon les fonctions et les rôles réservés aux évaluateurs de chaque type d'évaluation ou selon qu'ils soient évaluateurs internes ou externes.

L'ensemble de l'information provenant des diverses sources a permis de construire un profil comportant 40 compétences regroupées en sept dimensions. Ces compétences sont elles-mêmes associées à 153 tâches

différentes. À ces dimensions, compétences et tâches du profil s'ajoutent 34 contenus de formation et 40 caractéristiques ou qualités attendues d'une ou d'un évaluateur de programmes retenus comme pouvant faire partie du profil de compétences.

### La consultation des experts pour la complémentarité du profil

Le profil théorique a été soumis à 5 experts œuvrant dans la fonction publique qui ont accepté de compléter et de commenter la version théorique du profil au printemps 2004. Les commentaires recueillis ont permis de bonifier et de produire une version améliorée qui fut examinée par les 20 autres experts durant l'hiver 2005. En tout, 7 directeurs, 3 coordonnateurs et 15 professionnels de l'évaluation ont répondu au

questionnaire qui leur a été transmis. Au terme de cette consultation, des modifications ont été apportées à la première version du profil de compétences, la majorité d'entre-elles concernant la précision ou la simplification du langage utilisé ou du libellé des dimensions, des compétences ou des tâches ou encore leur ordre de présentation.

Cette activité de consultation et d'expertise correspondait dans les faits aux cinq premières étapes du modèle d'analyse: 1) l'identification des compétences de l'évaluateur, 2) la consultation des experts à titre de conseiller et d'expert professionnel, 3) la préparation du profil de compétences et du tableau des spécifications, 4) la constitution de la banque de tâches, et 5) l'identification des facteurs tâches. Les experts ont convenu de retenir sept dimensions et 47 compétences.

**TABLEAU 2 : LES SEPT DIMENSIONS ET LES 47 COMPÉTENCES DU PROFIL**

DIMENSIONS	COMPÉTENCES
1. Identification et analyse des besoins prioritaires	<ul style="list-style-type: none"> <li>· Repérer des besoins non comblés</li> <li>· Préciser les besoins non comblés</li> <li>· Déterminer la priorité des besoins</li> <li>· Communiquer adéquatement les besoins prioritaires au demandeur</li> </ul>
2. Construction de modèles logiques et formulation d'objectifs	<ul style="list-style-type: none"> <li>· Conceptualiser la logique d'un programme</li> <li>· Énoncer les normes de rendement ou les critères de succès d'un programme</li> <li>· Décrire le contexte d'un programme</li> <li>· Déterminer le bien-fondé de l'évaluation</li> </ul>
3. Confection du devis d'évaluation	<ul style="list-style-type: none"> <li>· Choisir l'école de pensée ou l'approche pertinente</li> <li>· Décider du type d'évaluation</li> <li>· Élaborer des stratégies adéquates</li> <li>· Rédiger, dans un format convivial, un devis, un cadre, une étude préparatoire, etc.</li> </ul>
4. Réalisation d'une évaluation de programme	<ul style="list-style-type: none"> <li>· Décrire la mise en œuvre du programme</li> <li>· Analyser correctement la situation dans laquelle évolue le programme en évaluation</li> <li>· Colliger les données requises à l'évaluation</li> <li>· Comparer les données réelles avec celles planifiées dans le plan du programme ou avec des normes de rendement</li> <li>· Expliquer les raisons pour lesquelles la mise en œuvre s'est éloignée du plan</li> <li>· Distinguer les « extrants » et les effets d'un programme</li> <li>· Décrire les « extrants » et les effets d'un programme</li> <li>· Reconnaître les possibilités d'effets inattendus ou non désirés</li> <li>· Associer les effets à des interventions précises</li> <li>· Déceler les facteurs qui affectent l'efficacité</li> <li>· Utiliser les modèles d'évaluation du rendement</li> <li>· Déterminer le mérite ou la valeur du programme</li> <li>· Utiliser les modèles théoriques appropriés et leurs cadres conceptuels</li> <li>· Utiliser les stratégies d'enquête quantitatives et qualitatives</li> <li>· S'assurer de la fidélité et de la validité des mesures</li> <li>· Définir les critères ou les normes d'interprétation</li> <li>· Interpréter les résultats de l'évaluation</li> </ul>



5. Administration et gestion de l'opération globale de l'évaluation	<ul style="list-style-type: none"> <li>· Estimer les coûts d'une évaluation</li> <li>· Planifier les ressources requises pour la réalisation d'une évaluation</li> <li>· Assurer l'efficacité des ressources évaluatives</li> <li>· Intégrer les résultats dans le processus de reddition de comptes</li> </ul>
6. Gestion des ressources humaines	<ul style="list-style-type: none"> <li>· Susciter la participation des divers intervenants</li> <li>· Intégrer l'évaluation aux pratiques des intervenants ou à la culture du programme</li> <li>· Consulter les intervenants</li> <li>· Exercer un leadership dans la gestion des situations difficiles</li> <li>· Partager les résultats avec les intervenants</li> <li>· Entretenir des relations positives avec les intervenants</li> <li>· Générer de nouveaux partenariats</li> <li>· Assurer la formation en évaluation des gestionnaires du programme et des intervenants</li> </ul>
7. Expertise-conseil en évaluation de programmes	<ul style="list-style-type: none"> <li>· Utiliser les principaux outils</li> <li>· Se tenir informé des nouvelles recherches et des nouvelles pratiques</li> <li>· Suggérer des pratiques adaptées aux situations et aux grands courants du domaine</li> <li>· Dégager les principaux constats</li> <li>· Faire des recommandations justes et réalistes à la suite d'une démarche d'évaluation</li> <li>· Assurer le suivi de la réalisation des recommandations issues d'un rapport d'évaluation</li> </ul>

### La validation du profil par des spécialistes de l'évaluation de programmes au Québec

Un questionnaire électronique autoadministré fut transmis aux membres de la Société québécoise d'évaluation de programmes (SQEP). Spécialistes en évaluation de programmes ou simplement intéressés ou encore actifs dans le domaine, 55 d'entre eux se sont prononcés sur la pertinence du contenu, sur l'adéquation des dimensions et des tâches, sur leur importance relative et sur l'exhaustivité du profil.

Parmi les répondants, 20 % occupent des fonctions d'encadrement (cinq directeurs et six coordonnateurs) alors que la moitié agissent à titre de professionnels et près de 75 % d'entre eux occupent des postes permanents. Ils sont, dans plus de 80 % des cas, employés d'organismes publics ou parapublics.

L'évaluation de programmes occupe plus de la moitié des fonctions accomplies par près de 65 % d'entre eux. La totalité des répondants ont une formation universitaire et 85 % d'entre eux possèdent un diplôme de deuxième ou de troisième cycle (40 maîtrises et 7 doctorats) dont le quart dans le domaine de la mesure et de l'évaluation.

Les opinions des 55 répondants de la SQEP ont permis de construire un profil en sept dimensions et 47 compétences que l'on a pu regrouper selon leur valeur essentielle. Le tableau 3 permet de visualiser ce profil dans ces diverses composantes et niveaux. Il s'agit en fait de notre profil de compétences des agents de recherche et de planification socioéconomique œuvrant à titre de spécialiste en évaluation de programmes.

**TABLEAU 3 : LE PROFIL DE COMPÉTENCES DITES DE BASE POUR OCCUPER UN POSTE EN ÉVALUATION DE PROGRAMMES À LA FONCTION PUBLIQUE QUÉBÉCOISE**

DIMENSIONS	PREMIER GROUPE	DEUXIÈME GROUPE	TROISIÈME GROUPE	QUATRIÈME GROUPE	AUTRES GROUPE
<b>Dimension 4.</b> Réalisation d'une évaluation de programme	<b>4.17</b> Interpréter les résultats de l'évaluation	<b>4.15</b> S'assurer de la fidélité et de la validité des mesures	<b>4.3</b> Colliger les données requises à l'évaluation <b>4.4</b> Comparer les données réelles avec celles du plan du programme ou avec des normes de rendement <b>4.5</b> Expliquer les raisons pour lesquelles la mise en œuvre s'est éloignée du plan <b>4.6</b> Distinguer les « extrants » et les effets d'un programme <b>4.7</b> Décrire les « extrants » et les effets d'un programme <b>4.10</b> Repérer les facteurs qui affectent l'efficacité <b>4.14</b> Utiliser les stratégies d'enquêtes quantitatives et qualitatives	<b>4.2</b> Analyser correctement la situation dans laquelle évolue le programme <b>4.8</b> Reconnaître les possibilités d'effets inattendus ou non désirés	<b>4.16</b> Définir les critères ou les normes d'interprétation <b>4.1</b> Décrire la mise en œuvre du programme <b>4.9</b> Associer les effets à des interventions précises <b>4.13</b> Utiliser les modèles théoriques appropriés et leurs cadres conceptuels <b>4.12</b> Déterminer le mérite ou la valeur du programme <b>4.11</b> Utiliser les modèles d'évaluation du rendement
<b>Dimension 2.</b> Construction de modèles logiques et formulation d'objectifs	<b>2.1</b> Conceptualiser la logique d'un programme		<b>2.2</b> Énoncer les normes de rendement ou les critères de succès d'un programme	<b>2.3</b> Décrire le contexte d'un programme	<b>2.4</b> Déterminer le bien-fondé de l'évaluation
<b>Dimension 3.</b> Confection de devis d'évaluation		<b>3.3</b> Élaborer des stratégies adéquates	<b>3.2</b> Décider du type d'évaluation	<b>3.4</b> Rédiger, dans un format convivial, un devis, un cadre, une étude préparatoire, etc.	<b>3.1</b> Choisir l'école de pensée ou l'approche pertinente
			<b>1.3</b> Déterminer la priorité des besoins <b>1.4</b> Communiquer adéquatement les besoins prioritaires au demandeur		<b>1.1</b> Déceler les besoins non comblés <b>1.2</b> Préciser les besoins non comblés
<b>Dimension 7.</b> Expertise-conseil en évaluation de programmes	<b>7.5</b> Faire des recommandations justes et réalistes à la suite d'une démarche d'évaluation	<b>7.4</b> Dégager les principaux constats		<b>7.3</b> Suggérer des pratiques adaptées aux situations et aux grands courants du domaine <b>7.6</b> Assurer le suivi de la réalisation des recommandations issues d'un rapport d'évaluation	<b>7.1</b> Utiliser les principaux outils <b>7.2</b> Se tenir informé des nouvelles recherches et des nouvelles pratiques

DIMENSIONS	PREMIER GROUPE	DEUXIÈME GROUPE	TROISIÈME GROUPE	QUATRIÈME GROUPE	AUTRES GROUPES
<b>Dimension 5.</b> Administration et gestion de l'opération globale de l'évaluation		<b>5.3</b> Assurer l'efficacité des ressources évaluatives		<b>5.2</b> Planifier les ressources requises pour la réalisation d'une évaluation <b>5.4</b> Intégrer les résultats dans le processus de reddition de comptes	<b>5.1</b> Estimer les coûts d'une évaluation
<b>Dimension 6.</b> Gestion des ressources humaines	<b>6.5</b> Partager les résultats avec les intervenants		<b>6.1</b> Susciter la participation des divers intervenants <b>6.2</b> Intégrer l'évaluation aux pratiques des intervenants ou à la culture du programme <b>6.3</b> Consulter les intervenants <b>6.6</b> Entretenir des relations positives avec les intervenants		<b>6.4</b> Exercer un leadership dans la gestion des situations difficiles <b>6.7</b> Gérer de nouveaux partenariats <b>6.8</b> Assurer la formation en évaluation des gestionnaires du programme et des intervenants

La composition et l'organisation de ce profil confirment les conclusions de Stevahn et al. (2005) pour lesquels, un tel profil spécifique et multidimensionnel avec des compétences hiérarchisées, existe bien.

Ces auteurs souhaitaient une validation du modèle. Nous croyons y être arrivés même si la structure du profil que nous présentons est plus élaborée en ce sens que des nuances de niveaux ont pu être apportés entre les dimensions et entre les compétences. Ce profil contient cinq dimensions certaines, mais plus probablement 7, 23 compétences évidentes et 24 autres fortement probables.

### ■ EN CONCLUSION

Le spécialiste en évaluation de programmes est un universitaire diplômé de deuxième cycle (maîtrise) ou de troisième cycle (doctorat) le plus souvent en sciences humaines ou en sciences de l'administration (mesure et évaluation de programmes). Il occupe généralement un emploi à plein temps au sein de la fonction publique québécoise ou dans une organisation paragonnementale.

Les principales tâches qu'il assume sont la réalisation d'évaluations, la construction de modèles logiques, la formulation d'objectifs et la confection de devis d'évaluation. On attend aussi de lui qu'il détecte et analyse les besoins prioritaires et fournisse une expertise-conseil en évaluation de programmes.

Parmi les compétences dont il doit faire preuve, certaines sont fondamentales et absolument essentielles. Ainsi doit-il interpréter les résultats de l'évaluation, s'assurer de la fidélité et de la validité des mesures, conceptualiser la logique d'un programme, concevoir des stratégies adéquates, faire des recommandations justes et réalistes à la suite d'une démarche d'évaluation, dégager les principaux constats, assurer l'efficacité des ressources évaluatives et partager les résultats avec les intervenants.

Un bon évaluateur possède une excellente capacité d'analyse. Il est méthodique et démontre un talent de communicateur. Il possède nécessairement des connaissances avancées en méthodes de recherche et d'échantillonnage, en technique de collecte de données ainsi qu'en statistiques.

Le profil des compétences révèle son caractère multidimensionnel avec des interrelations et une hiérarchie qui lui confèrent un caractère unique, suffisant semble-t-il pour le juger spécifique à la profession d'évaluateur de programmes. Les constats de l'analyse des corrélations de même que ceux de l'analyse factorielle corroborent ces propos.

Le profil de compétence des évaluateurs de programmes est suffisamment spécifique, avec des habiletés qui lui sont propres, pour justifier éventuellement une voie de recrutement réservée au cas où ce serait nécessaire. Cependant, la présence dans ce profil de compétences qui ne sont pas exclusives à ce métier ou profession laisse supposer une sorte de partage transversal de certains savoirs, savoir-faire, savoir-être et « savoir faire faire » avec d'autres activités professionnelles du même ordre.

## BIBLIOGRAPHIE

- Beaujolin, F. (1999). « La gestion des compétences. Études de cas commentés », *Gestion des ressources humaines*, Cahiers 128, Paris, Nathan Université, 143 p.
- Berk, R. A. (1986). *Performance Assessment: Methods and Applications*, The Johns Hopkins University Press, 544 p.
- Bernier, J.-J., B. Pietrulewicz (1997). *La psychométrie : traité de mesure appliquée*, Québec, Gaëtan Morin Éditeur, 444 p.
- Bertrand, O. (1997). *Évaluation et certification des compétences et qualifications professionnelles*, Programme de recherche et d'études de l'IPE, Institut international de planification de l'éducation, Paris, France, 128 p.
- Bloom, B. S., J. T. Hastings, G. F. Madeaus (1971). *Handbook on Formative and Summative Evaluation of Student Learning*, McGraw-hill, 923 p.
- Brien, R. (1997). *Science cognitive et formation*. Presses de l'Université du Québec, Sainte-Foy.
- Brochier, D. (2002). *La gestion des compétences : acteurs et pratiques*, Connaissance de la gestion, Paris, Économica, 179 p.
- Bordeleau, Y. (1992). *Gestionnaires performants. Comment expliquent-ils leur succès ?*, Éditions Agence D'Arc, 217 p.
- Bourgeault, J., M. Charih, D. Maltais, L. Rouillard (2003). *Les rôles et les compétences des gestionnaires supérieurs au Gouvernement du Québec*. Centre d'expertise en Gestion des Ressources Humaines SCT, Gouvernement du Québec. (Version électronique disponible sur le site du SCT), 50 p.
- Bujold, C., M. Gingras (2000). *Choix professionnel et développement de carrière. Théories et recherches*, 2<sup>e</sup> édition, Boucherville, Gaëtan Morin Éditeur, 421 p.
- Cook, T. D., L. C. Leviton, W. R. Shadish Jr. (1994). Program Evaluation. In the *Handbook of Social Psychology*, 432 p.
- Côté, L. (1993). Rapport de la Commission concernant les besoins de la relève des gestionnaires pour la fonction publique québécoise. *Sources-ENAP*, vol. 9, n° 5.
- Denimal, P. (2004). *Classification, qualification, compétences. Pour des actions sur l'organisation et le dialogue social*, France, Éditions liaisons, 254 p.
- Deshaies, P. et al. (1996). « Processus de planification d'un cours centré sur le développement d'une compétence », *Le Pôle de l'Est*. Sherbrooke, Performa.
- De Landsheere, V. (1988). *Faire réussir, faire échouer : la compétence minimale et son évaluation*, Presses Universitaires de France, 255 p.
- Dion, G. (1996). *Dictionnaire canadien des relations du travail*, 2<sup>e</sup> édition, Les Presses de l'Université Laval, 993 p.
- Dionne, P. (2004). *Prendre le leadership des compétences. Le réveil du management*, ERPI, Saint-Laurent, Québec, 161 p.
- Edwards, A. (1970). *The Measurement of Personality Trait*. Holt, Rinehart and Winston, inc. University of Washington, 306 p.
- Emploi-Québec (1993). *Guide de production d'une analyse de métier ou de profession*, Direction de l'Intervention

- sectorielle et territoriale, Gouvernement du Québec.
- Ferguson, G. A. (1971). *Statistical Analysis in Psychology and Education*, 3<sup>e</sup> édition, New-York, McGraw-Hill Book Company, 492 p.
- Fine, S. A. (1986). Job Analysis. Dans Berk, R.A. (1986) *Performance Assessment: Methods and Applications*. The Johns Hopkins University Press, pp. 53-81.
- Gale, L. (1998). *Discover What You're Best At*. The National Career Aptitude System and Career Directory. Simon and Schuster, Fireside, New York, 181 p.
- Gilbert, P., G. Schmidt (1999). *Évaluation des compétences et situations de gestion*, Connaissance de la gestion, Paris, Économica, 161 p.
- Guay, M.-M. (1998). Le profil de compétences en gestion : un instrument de rétroaction à 360 degrés pour le développement des cadres. Conférence prononcée au 10<sup>e</sup> Congrès international de l'Association internationale de psychologie du travail de langue française, à l'Université de Bordeaux.
- Guay, M.-M. (1997). *Performance et secteur public : réalités, enjeux et paradoxes*, Presses de l'Université du Québec, 397 p.
- Guba, E. G., Y. S. Lincoln (1981). *Effective Evaluation*, Jossey-Bass, 423 p.
- Gouvernement du Canada (2004). *Profil de compétences en formation et en apprentissage pour les conseillers en organisation et en classification*, Ottawa, Canada, 29 p.
- Herman, J. L., C. Taylor Fitz-Gibbon, L. Lyons Morris (1990). *Evaluator's Handbook*. Program Evaluation Kit, 2<sup>e</sup> édition, Sage Publications, 159 p.
- Joras, M. (2002). *Le bilan de compétences*, Que sais-je, PUF, France, 127 p.
- Kaplan, R. M., D. P. Saccuzzo (2005). *Psychological Testing. Principles, Applications, and Issues*, 6<sup>e</sup> édition, Thomson and Wadsworth, 745 p.
- Kerlinger, F. N. (1986). *Foundations of Behavioral Research*, 3<sup>e</sup> édition, Holt, Rinehart and Winston Inc, 667 p.
- Korenblit, P., C. Nicholas, H. Lehongre (2003). *Construire son projet professionnel à partir du bilan de compétences*, Collection Formation permanente, ESF éditeur, France, 224 p.
- Lasnier, F. (2000). *Réussir la formation par compétences*, Montréal, Guérin, 485 p.
- Le Boterf, G. (2002). *Ingénierie et évaluation des compétences*, 4<sup>e</sup> édition, Paris, Éditions d'organisation, 562 p.
- Legendre, R. (1993). *Dictionnaire actuel de l'éducation*, Montréal, Guérin, 1500 p.
- Marcelin, O., D. Bouteiller (1996). « Profils de compétence et performance des équipes de direction : quatre études de cas dans le réseau hospitalier québécois », dans A. M. Fericelli, B. Sire (1996), *Performance et ressources humaines*. Économica, 335 p.
- Maltais, D., B. Mazouz (2004). « À nouvelle gouvernance, nouvelles compétences : les compétences-clefs des gestionnaires de demain », *Revue Gestion*, vol. 29, n° 3, p. 82-92.
- Morin, E. M. (1996). *Psychologies au travail*, 2<sup>e</sup> édition, Boucherville, Gaëtan Morin Éditeur, 535 p.
- OCDE (2004). *Orientation professionnelle et politique publique. Comment combler l'écart*, Éducation et culture, Communauté européenne, France, 187 p.
- OCDE (2004). *L'orientation professionnelle : guide pratique pour les décideurs*, Éducation et culture, Communauté européenne, France, 82 p.
- Patton, M. Q. (1990). *Qualitative Evaluation and Research Methods*, Newbury Park, Sage, 532 p.
- Payette, A. (1988). *L'efficacité des gestionnaires et des organisations*, Sainte-Foy, Presses de l'Université du Québec, 310 p.
- Payette, A. (1992). « Les compétences des managers publics », dans R. Parenteau, *Management public : Comprendre et gérer les institutions de l'État*, Presses de l'Université du Québec, 640 p.
- Payette, A. (1997). « Les compétences des gestionnaires et la performance des organisations : essai de synthèse » dans M.-M. Guay (dir.), *Performance et secteur public – réalités, enjeux et paradoxes*, Sainte-Foy, Presses de l'Université du Québec, p. 7-30.

- Pemartin, D. (1999). *Gérer par les compétences ou comment réussir autrement ? L'actualité de la gestion*, Paris, Édition Management Société, 126 p.
- Perrenoud, P. (1995). Des savoirs aux compétences. De quoi parle-t-on en parlant de compétence? *Pédagogie collégiale*, 9 (1), pp. 20-24.
- Perrenoud, P. (1999). Construire des compétences, est-ce tourner le dos aux savoirs? *Pédagogie collégiale*, 12 (3), pp. 14-17.
- Popham, W. J. (1975). *Educational Evaluation*, Prentice-Hall, 328 p.
- Reuchlin, M. (1964). *Méthodes d'analyses factorielles*, École Pratique des Hautes Études, Bibliothèque scientifique internationale, Psychologie, Presses Universitaires de France, 418 p.
- Roegiers (2000). *Une pédagogie de l'intégration: compétences et intégration des acquis dans l'enseignement*, Bruxelles, De Boeck Université, 364 p.
- Rossi, P., H. Freeman, M. W. Lipsey (1999). *Evaluation: A Systematic approach*, 6<sup>e</sup> édition, Sage Publications, 500 p.
- Lecompte, R., L. Ruthman (1982). *Introduction aux méthodes de recherche évaluative*, Presses de l'Université Laval, 187 p.
- Santelmann, P. (2002). *Qualifications ou compétences. En finir avec la notion d'emplois non qualifiés*, Éditions liaisons, France, 244 p.
- Scallon, G. (2000). Notes de cours sur intranet. Université Laval.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*, L'école en mouvement, ERPI, Québec, 342 p.
- Scriven, M. (1978). *Merit vs Value*, *Evaluation News*, 8, 20-29.
- Stake, R. E. (1976). *Evaluating Educational Programs: The Need and the Response*, OECD Publications Center, 101 p.
- Stevahn, L., J. A. King, G. Ghere, J. Minnema (2005). Establishing Essential Competencies for Program Evaluators. *American Journal of Evaluation*, vol. 26, n° 1, mars, p. 43-59.
- Storey J. (2004). *Leadership in Organizations: Currents Issues and Key Trends*, Routledge, Taylor and Francis Group, London et New York, 349 p.
- Stufflebeam D. L. (1980). *L'évaluation en éducation et la prise de décision*, Victoriaville, Éditions HHP, 464 p.
- Tourmen, C., M. Basle et al. (2005). *Éléments pour la construction d'un référentiel des métiers de l'évaluation des politiques publiques*, Groupe de travail mandaté par la Société française d'Évaluation et l'European Evaluation Society, Université de Rennes I, France, 29 p.
- Toussaint, R. M. J., C. Xypas, (2004). *La notion de compétence en éducation et en formation. Fonctions et enjeux*, L'Harmattan, Paris, France, 319 p.
- Tresanini, M. (2004). *Évaluer les compétences: du recrutement à la gestion de carrière*, Éditions Management Société, collection Pratiques d'Entreprises, France, 275 p.
- Tyler, R.W. (1950). *Basic Principles of Curriculum and Instruction*, University of Chicago Press, 83 p.
- Tziner, A., C. Jeanrie, S. Cusson (1993). *La sélection du personnel: concepts et applications*, Montréal, Éditions Agence D'ARC, 227 p.
- Weiss, C. H. (1998). *Evaluation*, 2<sup>e</sup> édition, Prentice Hall, 372 p.
- Wolf, J.-L. (1998). *Méthodes de travail et stratégies d'apprentissage: du secondaire à l'université*, Bruxelles, De Boeck et Larcier, S.A, 325 p.
- Zorzi, R., M. McGuire, B. Perrin (2002). *Projet de la Société canadienne d'évaluation visant à appuyer la défense des intérêts et le perfectionnement professionnel. Avantages, extraits et éléments de connaissance de l'évaluation*, SCE, 78 p.

# L'utilité de l'évaluation : des conditions favorables et des pratiques variées

Par **André Viel**, Directeur de la planification et de l'évaluation, ministère du Développement économique, de l'Innovation et de l'Exportation (MDEIE) du Québec • [andre.viel@mdeie.gouv.qc.ca](mailto:andre.viel@mdeie.gouv.qc.ca)

Certaines personnes croient que la principale utilité d'une évaluation est de présenter les faits permettant de justifier l'abolition d'un programme. En réalité, les résultats des évaluations sont utilisés surtout pour améliorer les programmes en identifiant ce qui fonctionne bien, ce qui ne fonctionne pas et ce qui pourrait être fait pour bonifier le programme. Malgré cette approche constructive, il n'est pas exclu que le décideur puisse s'appuyer sur « ce qui ne fonctionne pas » pour décider de supprimer un programme ou un organisme. Nous verrons plusieurs exemples de la variété des pratiques dans l'utilisation de l'information contenue dans les rapports d'évaluation. Dans le cadre de mes fonctions d'administrateur public responsable d'unités d'évaluation, j'ai pu constater que certaine information de base, produite par les conseillers en évaluation, était beaucoup plus utile qu'on ne le croit.

Avant de présenter ces pratiques, un rappel des principales conditions favorisant l'utilisation des résultats des évaluations nous permettra de clarifier la notion d'utilité.

## ■ LES CONDITIONS FAVORISANT L'UTILISATION DES RÉSULTATS DES ÉVALUATIONS

Pour le *Center for Disease Control* ([www.cdc.gov/eval/standard.htm](http://www.cdc.gov/eval/standard.htm)), une évaluation sera d'autant plus utile qu'elle répond aux sept conditions suivantes : identification et implication des parties prenantes, crédibilité de l'évaluateur, étendue de l'information, présentation des critères d'interprétation des résultats, clarté du rapport, actualité et diffusion du rapport, suivi de l'évaluation.

## Identification et implication des parties prenantes

Plus il y a de parties prenantes impliquées dans l'évaluation, plus le nombre d'utilisateurs potentiels des résultats augmente.

Les premières personnes invitées à faire partie du comité d'évaluation d'un programme sont habituellement le gestionnaire du programme et le chargé de projet. Il s'agit en fait du cadre qui doit répondre de la gestion du programme et du professionnel qui assure le déroulement des étapes nécessaires au fonctionnement du programme. On peut s'attendre à ce que ces deux personnes jouent un rôle de premier plan dans la validation du devis d'évaluation et du rapport d'évaluation. Elles pourront faire valoir leurs besoins d'information et contribuer à préciser les questions d'évaluation. Elles seront aussi les premières à utiliser les résultats de l'évaluation et à se prononcer sur les recommandations.

La participation d'un bénéficiaire du programme au comité d'évaluation comporte plusieurs avantages : j'ai pu constater qu'elle permettait de rehausser le niveau des échanges, car la présence d'un participant externe peut prévenir les affrontements qui ont lieu parfois entre les personnes d'une même organisation. De plus, ce bénéficiaire du programme peut apporter une aide précieuse au moment de valider les questionnaires (compréhension des questions, accessibilité de l'information, etc.). L'utilité de l'évaluation pour ce client participant au comité d'évaluation est cependant limitée. Il pourrait en retirer certains avantages comme une meilleure compréhension du programme et des conditions d'admissibilité, mais ce n'est pas lui qui va utiliser les résultats de l'évaluation pour améliorer le programme ou décider de son avenir.

Le sous-ministre adjoint, de qui relève le programme, et le comité exécutif du ministère ne participent habituellement pas au comité d'évaluation, mais ils

devraient participer à la sélection des programmes à évaluer et être informés des résultats de l'évaluation au moment de la présentation du rapport, des recommandations et des suites envisagées par le gestionnaire. Leur implication à l'étape de l'approbation du cadre d'évaluation permet de valider le choix des questions et de prendre en compte leurs besoins. L'utilité des évaluations pour ces administrateurs de haut niveau ne doit pas être sous-estimée. J'ai pu constater le grand intérêt suscité par la présentation de certains rapports d'évaluation au comité exécutif, étant donné que les sous-ministres adjoints et le sous-ministre ont une connaissance parfois limitée des programmes de leurs ministères, surtout si ces programmes ne sont pas directement sous leur responsabilité. L'utilité est encore plus grande pour le sous-ministre et les sous-ministres adjoints qui entrent en fonction, une situation assez courante actuellement.

L'évaluation peut aussi être utile aux partenaires, qu'il s'agisse d'un autre ministère ou organisme participant au financement du programme, d'un représentant de réseau ou d'un expert. Leur participation au comité d'évaluation permet souvent d'éviter que le programme soit évalué plusieurs fois. Ce type de collaboration est hautement souhaitable compte tenu des ressources limitées en évaluation.

En résumé, l'évaluation peut être utile à plusieurs personnes et plus elles seront associées à l'évaluation, plus grandes seront les chances que l'information qu'elle contient soit utilisée.

### **Crédibilité de l'évaluateur**

Je ne crois pas que l'on puisse dire que plus un évaluateur est crédible, plus l'évaluation sera utilisée. D'autres conditions mentionnées dans cet article me semblent plus importantes. Par contre, si l'évaluateur n'est pas crédible, l'utilisation de l'évaluation peut être compromise. Il faut donc s'assurer de la crédibilité de l'évaluateur: il doit être compétent et digne de confiance, car l'acceptation des résultats en dépend.

Avoir une maîtrise en évaluation ou avoir plusieurs années d'expérience en évaluation constituent des qualifications contribuant à la crédibilité de l'évaluateur. Ces atouts peuvent servir à rassurer le client de l'évaluation au départ, mais ne suffisent pas à

assurer la crédibilité. Celle-ci repose sur les habiletés de l'évaluateur en communication écrite et orale, le réalisme de la stratégie d'évaluation qu'il propose et sa capacité de présenter des documents structurés, clairs, concis et surtout vulgarisés, sans erreur de calcul et sans faute.

Lorsque le milieu de travail est ouvert et favorise l'entraide, les évaluateurs évitent plusieurs erreurs et souvent imitent les pratiques des meilleurs. Il arrive aussi que le gestionnaire responsable de l'unité d'évaluation soit lui-même expérimenté et pose les bonnes questions ou effectue une révision serrée des documents, ou encore, dirige les réunions du comité d'évaluation. Toutes ces actions contribuent à rehausser la crédibilité de l'évaluateur, en particulier celui qui en est à ses premières armes en évaluation.

### **Étendue de l'information**

Lorsque l'évaluation repose sur l'analyse d'information pertinente, provenant de sources variées et étendues, son utilité a toutes les chances d'augmenter. Une évaluation utile est une évaluation qui nous apprend quelque chose sur la performance et la pertinence d'un programme, mais aussi sur des expériences similaires ou originales. Par contre, il est difficile de concilier les impératifs d'une évaluation qui doit être réalisée en moins de six mois et la nécessité de considérer un large bassin d'information. Une requête d'information auprès du gestionnaire du programme, une étude de la documentation accessible et un « sondage clientèle » permettent de constituer habituellement une base d'information intéressante. Mais un ou des groupes de discussion, des études de cas, des comparaisons avec d'autres juridictions ou des statistiques officielles nationales ou internationales augmentent l'utilité des évaluations; bien entendu, si le temps le permet.

Il faut résister à la tentation de s'en tenir à l'information accessible et de formuler les questions d'évaluation à partir de celle-ci. Renoncer à faire un sondage sous prétexte que l'information est accessible dans des fichiers ou dans des fiches de résultats, avant même d'avoir contrôlé la qualité de cette information, c'est risquer de baser l'évaluation sur des données incomplètes ou peu fiables.



### Présentation des critères d'interprétation des résultats

Pour moi, cela signifie qu'il faut annoncer d'avance à partir de quel point de vue, procédure ou raisonnement nous allons interpréter les résultats. Lorsque les attentes sont claires, il est plus facile de préciser le critère d'interprétation. Par exemple, si un organisme doit œuvrer auprès d'une clientèle constituée de PME, il devient possible d'énoncer que l'évaluateur s'attend à ce que celle-ci demeure stable ou augmente au cours de la période évaluée. Alors, si le résultat observé est qu'il y a migration au cours de la période vers une clientèle de plus en plus constituée de grandes entreprises ou d'entreprises publiques, l'organisme ne répond pas aux attentes.

La formulation d'hypothèses constitue aussi un autre moyen de spécifier les règles d'interprétation. Ainsi, plus il y a d'interventions auprès d'une entreprise visant l'adoption de meilleures pratiques d'affaires, plus sa productivité devrait augmenter.

Lorsque l'évaluation comporte un groupe contrôle, il est important qu'il y ait consensus sur la procédure pour le constituer sinon les résultats seront contestés sur la base des différences entre les membres du groupe. Il arrive que les clients de l'évaluation attendent les résultats pour décider s'ils contestent la composition du groupe contrôle.

À la suite de la présentation des résultats d'une évaluation, il se peut que le client de l'évaluation demande à l'évaluateur: «Finalement, compte tenu de tout ce qui a été dit dans le rapport, le programme est-il bon ou non?».

Évidemment, la question mérite d'être précisée: Le programme répond-il à un besoin? Est-ce que les objectifs sont atteints? Quels sont les effets du programme? Il y a plusieurs façons de vérifier si le programme répond à un besoin: l'évolution de la clientèle reflète habituellement la demande pour le programme; une augmentation de la clientèle peut signifier une croissance du besoin. Si les objectifs ne sont pas quantifiés, il sera plus difficile de savoir si le programme est bon du point de vue de l'atteinte des objectifs. Quant aux effets, il devient nécessaire de comparer avec des programmes similaires ou de suivre l'évolution des effets dans le temps.

### Clarté du rapport

Ici, on insiste sur la description claire du programme et de son contexte, et sur l'objet, la stratégie et les résultats de l'évaluation. J'ai pu constater que l'évaluation servait d'abord à augmenter la compréhension d'un programme: le modèle logique, la rétrospective sur trois ou cinq ans à l'aide des données sur les intrants, les activités et les produits et services, la validation des objectifs et de la rationalité du programme sont des moyens puissants qui permettent de saisir le sens de l'action du ministère en vue de changer les choses. Là où la clarté fait défaut, c'est souvent sur les effets nets. Mais là encore, il ne faut pas sous-estimer la capacité d'interprétation des clients des évaluations.

Le gouvernement intervient dans de nombreux domaines avec des moyens qui diminuent. On le constate en regardant fonder les crédits consacrés à certains programmes, ces crédits étant au départ le reflet d'une volonté de s'occuper du problème (le gouvernement n'est pas indifférent aux problèmes vécus par la population) bien plus que de le régler, étant donné l'écart entre les moyens et l'ampleur du problème. Malgré l'effet de levier des programmes qui entraîne la participation de partenaires privés ou publics, l'atteinte de la cible est souvent difficile à percevoir. Pourtant, l'effet sur les clients directs du programme est plus visible. À mon avis, les programmes ont une valeur ajoutée, mais elle est parfois faible, car certains bénéficient d'un financement très réduit reflétant le peu de moyens que le gouvernement peut leur consacrer.

### Actualité et diffusion du rapport

On comprend facilement qu'un rapport d'évaluation utile est un rapport qui arrive au bon moment et qui est diffusé aux bonnes personnes. Certaines dates de production d'un rapport sont incontournables: lorsqu'il s'agit de renouveler une convention de subvention qui vient à échéance ou lorsqu'il s'agit de produire une évaluation qui doit accompagner une demande de crédits budgétaires. Bien entendu, cette contrainte de temps conditionne grandement la profondeur de l'évaluation.

Par contre, il peut être bénéfique de faire coïncider la diffusion d'un rapport avec l'entrée en fonction

d'un nouveau sous-ministre adjoint ou d'un nouveau gestionnaire du programme.

Il importe que le rapport soit diffusé le plus largement possible, même si sa diffusion est soumise à certains délais.

### **Suivi de l'évaluation**

À ma connaissance, rares sont les unités d'évaluation qui font un suivi de leurs recommandations ou de leurs pistes d'amélioration. Savoir ce que le gestionnaire compte faire pour donner suite aux recommandations permet déjà d'avoir une bonne idée de l'utilité de l'évaluation. Plusieurs recommandations peuvent être suivies, d'autres rejetées. Mais parmi celles qui sont acceptées, combien sont réellement mises en œuvre? Pourtant, cette information est importante pour apprécier l'utilité de l'évaluation.

### **■ DES PRATIQUES VARIÉES DANS L'UTILISATION DES ÉVALUATIONS**

Nous venons de voir quelles sont les conditions pour qu'une évaluation soit utile. Dans les faits, on peut se demander à quoi ont servi les évaluations. Comme superviseur de nombreux rapports d'évaluation, j'ai pu constater la diversité des usages des rapports d'évaluation et l'utilité de l'information qu'ils contiennent.

### **Le renouvellement des conventions de subvention**

Les ministères et organismes gouvernementaux accordent des subventions à de multiples intervenants (centres de recherche, associations, organismes à but non lucratif, etc.) Habituellement, une convention précise les attentes envers ces intervenants et une clause de la convention prévoit la réalisation d'une évaluation avant la fin de la convention. Dans ce contexte, le rapport d'évaluation est utilisé pour apprécier la performance de l'organisme, pour vérifier s'il répond encore à un besoin et pour définir les attentes qui feront partie de la prochaine convention. Dans les faits, on retrouve parfois les pistes d'amélioration, identifiées à la suite de l'évaluation, reproduites telles quelles dans la convention de subvention.

Lorsque vient le temps de renouveler une convention de subvention, le gestionnaire aura plus de difficulté à spécifier les attentes en l'absence d'évaluation. De plus, peut-on de façon responsable renouveler la convention sans savoir ce qui s'est passé au cours de la période couverte par la convention précédente?

### **Les demandes de financement**

Le gestionnaire de programme ou l'organisme évalué peut se référer à l'évaluation pour obtenir un financement supplémentaire. Dans les faits, certains organismes n'ont pas hésité à citer le contenu de rapports d'évaluation dans le cadre de démarches de demande de financement et ils ont obtenu ce financement. Bien entendu, ces organismes ont fait ressortir les points positifs de l'évaluation à l'appui de leur démarche. Par contre, comme l'évaluation fait état des constats positifs et négatifs, celui qui reçoit la demande de financement peut toujours la consulter. L'évaluation permet ainsi de valider la crédibilité du demandeur qui pourrait pencher un peu trop du côté des faits qui appuient sa demande.

Il arrive aussi que le Secrétariat du Conseil du trésor ou le sous-ministre adjoint responsable du programme demande la réalisation d'une évaluation en vue d'éclairer la prise de décision dans le cadre du processus budgétaire. Pour sa part, le gestionnaire du programme, à qui on demande un mémoire pour appuyer une demande de financement, sera beaucoup mieux outillé si une évaluation a été réalisée. Sans évaluation, il devra rassembler en quelques semaines ou parfois en quelques jours, l'information qui lui permettra de préparer le mémoire, ce qui exclut la collecte de données originales sur la performance du programme. Un rapport d'évaluation constituera alors une source d'information très utile résultant d'un travail réalisé sur une période de plusieurs mois.

### **La connaissance des programmes**

Certains gestionnaires se servent des rapports d'évaluation pour faire connaître leur programme. À cet égard, les données sur l'évolution du financement du programme et de sa clientèle, le modèle logique et les comparaisons avec des programmes similaires fournissent un éclairage très apprécié. Certains gestionnaires ont même proposé de faire encadrer le

modèle logique tellement il facilitait la compréhension de la finalité du programme. Ils sont confiants que l'évaluation permettra d'établir la valeur de leur programme et ils utilisent les résultats de l'évaluation pour en faire la promotion auprès de leurs supérieurs immédiats et hiérarchiques, et aussi, auprès de leurs clients ou partenaires.

La composition des comités exécutifs des ministères change périodiquement. Les nouveaux venus et même les anciens ne connaissent pas toujours la réalité des programmes administrés par le ministère. La présentation des rapports d'évaluation au comité exécutif devient alors très utile. En augmentant la compréhension des programmes et de leurs résultats, l'évaluation peut prévenir les décisions à l'emporte-pièce ou les jugements basés sur des impressions. De plus, la présentation des rapports au comité exécutif permet de faire le point sur un programme et donne lieu parfois à des discussions prolongées ou à des demandes spécifiques qui s'adressent directement au gestionnaire du programme.

En diffusant les rapports aux clients et aux partenaires et en valorisant les « bons coups », le gestionnaire contribue à leur motivation et à la poursuite de leur engagement. Certains rapports sont peu diffusés, d'autres sont imprimés et distribués à des interlocuteurs ciblés, d'autres sont reproduits sur le site Internet du ministère. Les pratiques sont très variées à cet égard. Évidemment, pour favoriser l'utilité des évaluations, la plus large diffusion des rapports est souhaitable.

### Le suivi des programmes

Idéalement, les programmes devraient faire l'objet d'un suivi au moyen de fiches de résultats permettant d'apprécier la performance du programme. Dans les faits, les fiches de résultats comportent certaines limites : les gestionnaires du programme n'ont pas toujours le temps de s'assurer qu'elles sont produites et ils n'ont pas le temps non plus de les valider ; certains résultats se font jour longtemps après le versement de la tranche finale de la subvention. Malgré les avantages des fiches de résultats, je ne crois pas qu'elles puissent remplacer l'évaluation. En réalité, l'évaluation remplace souvent le suivi qui aurait dû être fait et qui ne l'a pas été, faute de temps. En réalisant une synthèse sur les principaux enjeux d'un

programme, l'évaluation produit une information plus complète que celle fournie par les tableaux de bord ou les fiches de résultats.

### Les effets des programmes

Sauf exception, l'évaluation telle que nous la pratiquons ne nous renseigne pas de façon satisfaisante sur les effets nets des programmes. On mesure assez bien la quantité et la qualité des produits et services livrés, de même que les bénéfices pour les clients du programme, mais les effets sur la cible sont difficiles à isoler. En présentant l'évolution des ressources consacrées au programme et l'effet de levier (montant investi par les partenaires / montant investi par le ministère), l'évaluation permet d'apprécier l'effort gouvernemental qui, devant l'ampleur de certains problèmes, demeure limité. Avec les moyens mis en œuvre, pouvons-nous vraiment changer la situation ?

Il se peut aussi que les évaluations servent à atténuer des conflits à l'intérieur de l'organisation. Il arrive que certains gestionnaires attribuent à leur programme des effets sans commune mesure avec la réalité, ce qui irrite au plus haut point certains collègues qui ont à partager le budget du ministère. Une évaluation qui vient remettre les résultats en perspective peut servir de monnaie d'échange : « Tu arrêtes de claironner tes résultats mirobolants, j'arrête de claironner les résultats de l'évaluation. »

### Abolir ou non un programme

Quel est le pourcentage des évaluations qui a donné lieu à l'abolition d'un programme ? Cette question a été posée en commission parlementaire et reflète bien les attentes que certains peuvent avoir concernant l'utilité de l'évaluation de programmes. Il est possible qu'une évaluation contribue directement à l'abolition d'un programme, mais, apparemment, ce serait l'exception. En fait, l'évaluation viendrait appuyer une décision déjà envisagée concernant un programme dont la performance médiocre est connue ou qui ne répond plus à un besoin ; elle viendrait en quelque sorte donner le coup de grâce.

Cette perception de l'utilité de l'évaluation, c'est-à-dire que l'évaluation sert à abolir les programmes,

est plus répandue qu'on ne le croit. Les cadres et les professionnels qui gèrent des programmes sont parfois craintifs lorsqu'ils apprennent que leur programme va être évalué. J'ai même entendu, lors d'une rencontre préparatoire à une évaluation, des participants manifester leur crainte de perdre leur emploi.

L'information contenue dans un rapport d'évaluation peut servir à justifier la survie ou l'abolition d'un programme. Comme les rapports révèlent les points forts et les points faibles du programme, le décideur peut s'appuyer sur cette information pour justifier son action. Normalement, des pistes d'amélioration sont proposées à la suite de l'évaluation et le programme est maintenu en supposant que le gestionnaire tiendra compte de ces pistes. Rien n'empêche cependant un sous-ministre de faire référence à certains constats du rapport pour justifier l'abolition du programme dans un contexte de restriction budgétaire.

## ■ CONCLUSION

Je crois que les responsables de l'évaluation de programmes ont avantage à faire valoir la contribution de l'évaluation à l'amélioration des programmes, car il s'agit là de sa principale utilité. L'évaluateur doit rendre compte de la réalité du programme, sa

réponse à un besoin, l'atteinte de ses objectifs et ses effets et proposer des pistes d'amélioration. Le décideur aura toujours le choix, en considérant le contexte budgétaire, administratif et politique, de maintenir le programme en exigeant qu'on donne suite aux pistes d'amélioration ou de l'abolir. En fait, l'évaluation fournit l'alternative à l'abolition pure et simple d'un programme en proposant de le maintenir moyennant certaines améliorations. Le décideur peut, par contre, considérer que les améliorations suggérées sont trop importantes, que les résultats sont décevants, que les besoins diminuent et conclure à la nécessité d'abolir le programme. L'évaluation vise à soutenir la prise de décision; celle-ci appartient en définitive au décideur.

J'ai souvent entendu des gestionnaires à qui l'on remettait un cadre d'évaluation ou un rapport d'évaluation faire le commentaire suivant: « C'est la première fois qu'on a cette information. » L'évaluation leur apporte des renseignements utiles: pistes d'amélioration, points forts et points faibles, évolution du financement et de la clientèle, comparaisons avec des programmes similaires, synthèse, qualité et quantité des produits et services livrés, résultats et autres.

# Le Centre de recherche et d'expertise en évaluation

Le Centre de recherche et d'expertise en évaluation (**CREXE**) de l'**ENAP** du Québec est actif depuis le 1<sup>er</sup> juin 2005. Sa mission est d'assumer un leadership dans le développement, l'application et la promotion de l'évaluation, tant au niveau national qu'international, ainsi que dans la formation de chercheurs et d'analystes en évaluation.

Le **CREXE** s'appuie sur un personnel hautement qualifié et expérimenté. Il fait appel en outre à l'expertise des professeurs de l'**ENAP** et associe des professionnels et des étudiants à ses projets de recherche.

Le **CREXE** oriente ses interventions dans trois rayons d'activité :

## ÉVALUATION DE POLITIQUES ET DE PROGRAMMES

Le Groupe d'évaluation de politiques et de programmes (GÉVA) a pour mandat l'évaluation de politiques, de programmes, de lois et de projets publics. Le GÉVA est en mesure de proposer une démarche d'accompagnement en trois étapes: une étude préparatoire à l'évaluation, une évaluation formative à mi-parcours et une évaluation sommative en fin de projet.

## ÉTUDE SUR LES POLITIQUES PUBLIQUES ET LA SANTÉ

Le Groupe d'étude sur les politiques publiques et la santé (GEPPS) est un partenaire du CREXE. Ce groupe a pour objet l'étude du processus de production des politiques publiques, de leur émergence à leur évaluation, vue sous l'angle de la santé. Le GEPPS vise également à favoriser l'appropriation, par les acteurs concernés, des connaissances pertinentes pour l'élaboration de politiques favorables à la santé.

## MESURE DE LA PERFORMANCE DES ORGANISATIONS PUBLIQUES

Le Groupe de mesure de la performance des organisations publiques (MPOP) s'intéresse à la théorie et aux pratiques de la mesure de la performance des organisations publiques. Le MPOP se penche également sur la vérification de l'optimisation des ressources (VOR) et sur les systèmes d'information qui, en complément à l'évaluation, permettent la mesure de la performance organisationnelle.

**POUR INFORMATION :**  
CENTRE DE RECHERCHE ET D'EXPERTISE EN ÉVALUATION (CREXE)  
ENAP  
555, BOUL. CHAREST EST  
QUÉBEC (QUÉBEC) G1K 9E5  
TÉLÉPHONE : (418) 641-3000, POSTE 6324



[www.crexe.enap.ca](http://www.crexe.enap.ca)

# L'évaluation de l'Éducation au Danemark

Par **Vibeke Normann Andersen**, University of Southern Denmark, Syddansk Universitet, Département de science politique et de gestion publique, Institut for Statskundskab • vna@sam.sdu.dk

Traduit de l'anglais

Depuis le début des années 1990, le Danemark s'est prêté à plusieurs analyses comparatives internationales effectuées par l'Association internationale pour l'évaluation du rendement scolaire (IEA) et l'OCDE. Les résultats de la première épreuve internationale de compréhension de l'écrit à laquelle participait le Danemark ont montré que les élèves danois de neuf ans détenaient une faible aptitude à la lecture et manquaient d'assurance comparativement aux élèves des pays avec lesquels le Danemark se compare habituellement. En revanche, à partir de l'âge de 14 ans, l'aptitude à la lecture des élèves danois était identique à celle des élèves des autres pays et sur le plan de la sociabilité, les élèves danois figuraient parmi les meilleurs. En présence de ces résultats, le Danemark s'est penché sur la faible performance des élèves danois en éducation de base, les résultats ayant été jugés insuffisants. D'autres épreuves, dans les disciplines scientifiques par exemple, ont donné par la suite des résultats semblables.

Ces résultats ont suscité un vif débat chez les enseignants, les politiciens et le public en général. La piètre performance des élèves a été dénoncée comme étant un problème de nature politique. Ces réactions tiennent sans doute au fait que les Danois estiment qu'ils possèdent l'un des meilleurs systèmes d'éducation au monde (Norrild, 2004, 16). Cette vision est très probablement liée au fait que le Danemark figure parmi les pays qui investissent le plus dans l'éducation (OECD, 2005). L'un des systèmes d'éducation les plus coûteux du monde se devait donc d'afficher de meilleures performances. Selon les dernières statistiques de l'OCDE publiées en 2005, le Danemark compte

effectivement parmi les pays membres qui investissent le plus dans l'éducation, qu'il s'agisse des ressources investies par étudiant ou du pourcentage du produit national brut (PNB) consacré à l'éducation. Toutefois, l'écart entre les résultats escomptés en fonction des ressources consenties et les résultats réellement obtenus demeure important. Le Danemark se classe au troisième rang parmi les 30 pays membres de l'OCDE pour les ressources consacrées à l'éducation, exprimées en pourcentage du PNB, mais n'occupe que le douzième rang quant aux aptitudes en mathématiques des élèves de 15 ans (OECD, 2005). Le ministre de l'Éducation du Danemark n'a pas jugé les résultats satisfaisants ([www.uvm.dk/05/pm130905.htm?menuid=641015](http://www.uvm.dk/05/pm130905.htm?menuid=641015)).

Le système d'éducation danois a connu alors plusieurs réformes. L'évaluation a été considérée comme un moyen de s'attaquer aux problèmes de la performance en éducation de base. La création d'une culture axée sur les résultats et l'évaluation a constitué l'une des principales tâches du ministère de l'Éducation et des associations de ce secteur d'activité. Comparativement aux autres pays, le Danemark a instauré tardivement l'évaluation dans le secteur public. Cependant, depuis le début des années 1990, les initiatives se sont multipliées dans le secteur de l'éducation par rapport à d'autres secteurs (Hansen et Hansen, 2001, 160). Malgré les critiques qui lui reprochent de ne pas suffisamment y recourir, le gouvernement du Danemark mise désormais sur l'évaluation comme moyen d'accroître la qualité de l'éducation (The Danish Government, 2002). Les établissements d'enseignement sont, par la loi, soumis à l'évaluation. Les autorités danoises ont ainsi adopté une nouvelle façon de gouverner axée sur les résultats et l'évaluation (The Danish Government, 2002).

Le présent article porte sur trois réformes qui instaurent la pratique de l'évaluation dans le secteur de l'éducation au Danemark. Tout d'abord, la création d'un Institut national d'évaluation chargé d'effectuer

des évaluations dans toutes les disciplines de l'éducation. Ensuite, l'adoption par le Parlement danois d'une loi sur la transparence et l'ouverture en éducation, en vertu de laquelle tous les établissements d'enseignement (des écoles primaires aux universités) doivent publier leur performance sur Internet. Dans cette loi, la performance est définie comme résultant des moyennes générales de tous les examens finals et de toutes les évaluations effectuées à l'exception des évaluations internes qui portent sur les enseignants. Finalement, le ministère de l'Éducation du Danemark encourage les établissements d'enseignement à effectuer des autoévaluations afin de créer une culture axée sur les résultats et l'évaluation. Bien que la plupart de ces réformes visent l'éducation en général, l'attention du public et des politiques est principalement dirigée vers l'enseignement primaire, sujet abordé dans cet article.

### ■ L'ÉVALUATION DE LA POLITIQUE D'ÉDUCATION

Au Danemark, la durée de la scolarité obligatoire est de neuf ans. Bien que l'éducation soit obligatoire, la fréquentation scolaire ne l'est pas. L'éducation est considérée comme un « outil de promotion de l'égalité » (Royal Danish Ministry of Foreign Affairs, 2003 *Udenrigsministeriet*). L'ambition est de permettre au plus grand nombre de Danois de recevoir une éducation de qualité. L'objectif du gouvernement danois (2002, 11) est de créer un système d'éducation capable d'égaliser le meilleur système du monde d'après les évaluations comparatives. L'Institut danois d'évaluation est désigné explicitement comme un moyen de bâtir un système d'éducation performant (Royal Danish Ministry of Foreign Affairs, 2003).

En matière d'éducation, l'État, les comtés ou régions et les administrations locales se partagent les responsabilités. Elles varient selon le niveau. En ce qui concerne l'enseignement primaire, le ministère de l'Éducation fixe les objectifs, les principales compétences pédagogiques requises, les programmes d'études et l'organisation des examens. En même temps, une vaste décentralisation caractérise cette éducation de base. Les écoles primaires et les écoles du premier cycle d'enseignement secondaire sont administrées par les administrations locales qui en assurent la gestion économique. La responsabilité des écoles du deuxième cycle d'enseignement secondaire

est partagée. Les *grammar schools* (écoles d'enseignement classique) sont sous la responsabilité des comtés du Danemark. Les écoles de formation professionnelle de base sont des établissements indépendants soumis au contrôle du ministère de l'Éducation. L'importante réforme prévue en 2007 remplacera les 14 comtés du Danemark par 5 régions. Par voie de conséquence, les écoles ordinaires du deuxième cycle d'enseignement secondaire, à l'instar des écoles de formation professionnelle de base, deviendront des établissements indépendants supervisés par le ministère de l'Éducation.

La décentralisation des pouvoirs vers les administrations locales et les institutions indépendantes explique en partie pourquoi l'évaluation est devenue une priorité pour les autorités nationales. Selon les concepts traditionnels de la gouvernance, l'État a la capacité d'assurer la prestation efficace de l'éducation (Pierre et Peters, 2000). On peut toutefois se demander si les autorités nationales ont la capacité de diriger et de contrôler des secteurs de plus en plus décentralisés comme l'enseignement primaire. L'évaluation apparaît dès lors comme un mécanisme de régulation.

La tendance qui consiste dorénavant à mettre l'accent sur l'évaluation de l'éducation plutôt que sur la politique de l'éducation doit être resituée dans le contexte de la modernité réflexive (Beck, Giddens et Lash, 1994). La rationalité qui fonde la société moderne est remise en cause. Un des plus grands problèmes auxquels l'État est confronté sont les effets secondaires internes du manque de coordination. Les autorités nationales tentent donc d'y remédier en faisant appel à l'évaluation. Par nature, elle questionne les valeurs acquises. « L'évaluation est le processus qui détermine le mérite et la valeur des choses » [*traduction libre*] (Vedung, 1997, 2). C'est une évaluation rétrospective systématique de ce mérite et de cette valeur, c'est-à-dire de la performance et des résultats des interventions gouvernementales, qui joue un rôle sans cesse plus important (Vedung, 1997, 3). Les évaluations peuvent révéler les paradoxes de la notion d'efficacité et le manque de coordination. En cette ère de postmodernisme, l'évaluation, en tant que nouveau mécanisme de contrôle, représente plus qu'un besoin fonctionnel. C'est un reflet de l'état sociologique actuel et une tentative de contrôle social au regard des incertitudes et du risque (Jordana et Levi-Faur, 2004, 14). Ainsi,

l'autoréflexion et l'autoévaluation deviennent-ils des éléments culturels dominants (Dahler-Larsen, 1998).

La politique danoise en matière d'éducation a longtemps permis à chaque région et à leurs établissements associés, tels que les écoles primaires et les écoles du deuxième cycle d'enseignement secondaire, de conserver leur propre système d'administration et de gestion (Christensen, 2000, 2003). La réglementation nationale varie, elle, tel qu'il a été décrit précédemment, en fonction du degré de décentralisation, de l'échelon du contrôle, central, régional ou local et aussi de la tradition de participation des usagers des établissements. Chaque région s'inspire donc d'un héritage historique singulier, ce qui a pour effet de créer une diversité complexe dans le secteur de l'éducation (Christensen, 2000, 199, 202-5).

La structure du ministère de l'Éducation reflète également la segmentation de la politique en matière d'éducation et la prédominance des politiques sectorielles (Knudsen, 2001, 46, 47). L'accent mis sur l'évaluation et l'atteinte de résultats peut donc être considéré comme une tentative des pouvoirs publics pour instaurer une cohérence dans la gestion de l'éducation. Ainsi les mécanismes d'évaluation mis en place soumettent-ils tous les établissements d'enseignement à la même réglementation. Dans cette perspective, le type d'évaluation choisi revêt une grande importance puisqu'il aura un impact direct sur la politique du Danemark en matière d'éducation. Nous consacrons la prochaine partie aux divers types d'évaluation.

### ■ L'INSTITUT DANOIS D'ÉVALUATION

L'Institut danois d'évaluation a été créé en 1999, à la suite des résultats insuffisants obtenus par les élèves danois lors des tests internationaux comparatifs (Dahler-Larsen, 2004, 65). C'est un organisme indépendant financièrement, à la fois du ministère de l'Éducation et des établissements (Leeuw, 2003, 14), mais qui doit tenir compte des directives du ministère. Son rôle consiste à surveiller et à évaluer l'ensemble des secteurs éducatifs. Il a le droit et l'obligation de décider et d'effectuer des évaluations des écoles primaires, secondaires et de formation professionnelle de même que des établissements d'enseignement supérieur, des établissements d'éducation des adultes et des universités.

Les évaluations peuvent être commandées à la demande du ministère, des commissions consultatives, des administrations locales ou des établissements d'enseignement. Les établissements d'enseignement n'ont d'autre choix que d'y participer (Leeuw, 2003, 14). Les évaluations sont conduites dans les établissements d'enseignement publics aussi bien que dans les établissements d'enseignement privés qui reçoivent des subventions publiques. Ces subventions représentent une part substantielle du budget de la plupart des établissements d'enseignement privés au Danemark. L'Institut procède de temps à autre à l'évaluation de la performance des établissements d'enseignement privés pour vérifier le bien-fondé des subventions.

Les évaluations effectuées par l'Institut peuvent donc avoir des répercussions importantes sur les activités et même la survie de certains établissements d'enseignement. Dans une note explicative sur sa loi constitutive, l'Institut danois d'évaluation est décrit en premier comme contribuant au développement et à l'assurance de la qualité de l'enseignement grâce à des évaluations indépendantes puis, ensuite, comme exerçant un contrôle réel de l'atteinte des objectifs de l'enseignement (Kristoffersen, 2003, 26). Nonobstant cette forte indépendance, les tâches de l'Institut danois d'évaluation diffèrent assez peu de celles dévolues aux organismes similaires dans d'autres pays malgré leur moindre degré d'indépendance (Leeuw, 2003, 15).

L'Institut danois d'évaluation a produit ce qu'il est convenu d'appeler le « modèle danois d'évaluation ». L'Institut souligne qu'il utilise les méthodes et pratique les types d'évaluation appropriés pour chaque discipline, par exemple pour l'évaluation, la vérification, la comparaison et l'accréditation d'un programme ou pour l'évaluation plus générale de l'enseignement ou des établissements (Kristoffersen, 2003). Quant aux évaluations réalisées, certaines s'attachent à comparer les établissements entre eux alors que d'autres étudient la transition entre le premier et le deuxième cycle d'enseignement secondaire.

Pour autant, le modèle danois d'évaluation prévoit des procédures communes à toutes les évaluations de l'Institut quel que soit leur contenu ou le type d'établissement évalué. Il existe donc bien une sorte d'uniformisation de l'évaluation. Au début du projet, une équipe est formée et des études initiales sont



menées. C'est également à cette étape que sont déterminées les modalités de l'évaluation, telles que l'objectif et le cadre de l'évaluation. Ensuite, un groupe d'évaluation externe composé d'universitaires spécialisés en évaluation est réuni. Finalement, une autoévaluation est effectuée par les membres de l'établissement visé. L'obligation d'effectuer une autoévaluation figure dans une note explicative sur la loi sur l'Institut danois d'évaluation (BEK n° 1117 of 28.12.2001). Cette autoévaluation doit suivre les lignes directrices arrêtées par l'Institut (Kristoffersen, 2003, 31). L'Institut peut également effectuer des enquêtes supplémentaires suivies de la visite des lieux par l'équipe de projet et le groupe d'évaluation. Enfin, l'Institut rédige un rapport qui précise à quelles fins est destinée l'évaluation.

Tous les rapports publiés par l'Institut doivent être mis en ligne sur son site Internet. De plus, les établissements qui ont fait l'objet d'une évaluation sont tenus d'élaborer un plan d'action indiquant de quelle façon ils prévoient de se conformer aux recommandations de l'Institut. Encore une fois, les règlements exigent que les établissements évalués rendent leur plan public en utilisant un moyen électronique (BEK n° 1117, § 4, part 2 of 28.12.2001). Compte tenu de son fondement législatif, des ressources et de l'organisation dont il dispose pour élaborer et mettre en place un modèle d'évaluation, l'Institut danois d'évaluation a un impact considérable sur la nature de l'évaluation dans le secteur de l'éducation. L'Institut doit aussi se concerter avec des groupes puissants et des représentants de la société civile. Les ministères, les organisations syndicales et patronales, les pouvoirs publics locaux et régionaux, les associations d'enseignants, etc. sont tentés d'intervenir dans les processus d'évaluation (Leeuw, 2003, 21) et d'influencer le modèle d'évaluation ou une évaluation en particulier.

### ■ TRANSPARENCE ET OUVERTURE EN ÉDUCATION

La loi sur l'Institut danois d'évaluation précisait que les rapports produits par l'Institut ne devaient pas faire référence aux notes ou aux résultats scolaires ni dresser un classement des établissements évalués. Cependant, à la suite d'une décision parlementaire prise en 2002 dans le cadre de la discussion de la loi sur la transparence et l'ouverture en éducation (Loi 414), la restriction au sujet des notes a été abolie.

La loi a introduit le contrôle de la performance dans le système d'éducation danois. Afin d'assurer une performance de haute qualité, tous les établissements d'enseignement sont tenus de documenter leur performance en publiant sur Internet leur énoncé de valeurs, leur philosophie pédagogique, leurs moyennes générales ainsi que les évaluations de la qualité de l'enseignement. Cette loi a donc imposé des changements substantiels dans le système d'éducation danois.

D'une part, les moyennes générales des établissements d'enseignement n'étaient pas accessibles auparavant. D'autre part, cette loi « oblige » les établissements d'enseignement à rédiger et à rendre public un énoncé de valeurs et une philosophie pédagogique. Enfin, chaque établissement d'enseignement doit, le cas échéant, créer sa propre page d'accueil. Ces mesures, avec la publication des résultats des évaluations, sont censées aider à faire remonter la côte de l'éducation danoise dans les comparatifs internationaux.

La raison d'être officielle de la loi sur la transparence et l'ouverture en éducation est de permettre aux établissements d'enseignement de partager leurs expériences et de tirer des leçons des « bonnes pratiques ». La communication de l'information évaluative sur la performance a pour but d'encourager les utilisateurs et les professionnels des établissements d'enseignement à travailler ensemble. En outre, toujours selon la raison d'être officielle, cette sensibilisation à la bonne performance sera renforcée grâce à la compétition entre les établissements. Le terme « compétition » entre les établissements publics ne fait généralement pas partie du discours gouvernemental au Danemark, mais il a été explicitement utilisé dans le commentaire officiel de la loi (L414). De surcroît, la population aura accès à de l'information sur la performance des écoles et des universités, ce qui lui permettra de choisir à bon escient parmi les divers établissements d'enseignement (Andersen et Dahler-Larsen, à paraître).

Publier l'information évaluative sur Internet est une fin en soi, mais également un moyen – espère-t-on – de se débarrasser de la note embarrassante que le Danemark a obtenu dans les évaluations comparatives internationales. La loi sur la transparence et l'ouverture met en exergue un certain nombre de concepts : la transparence, l'ouverture, l'assurance

de la qualité, le développement de la qualité et la compétition. Aucun modèle plus précis de comportement des utilisateurs, des enseignants et des chefs de file en éducation n'est cependant privilégié. (Andersen et Dahler-Larsen, à paraître).

Les énoncés de valeurs et les philosophies pédagogiques que chaque établissement d'enseignement se voit dans l'obligation d'élaborer, diffèrent des documents d'évaluation traditionnels, car ils sont prospectifs et ne nécessitent pas l'application d'une méthodologie précise. L'information prospective, combinée à l'information évaluative plus traditionnelle, permet en principe une meilleure approche de la performance de chaque établissement tout en leur assurant un certain degré d'autonomie (Andersen et Dahler-Larsen, à paraître). Avec la systématisation des évaluations et la non-publication des évaluations internes (évaluations d'enseignants ou d'employés) une ambiguïté subsiste cependant.

La loi fait référence en effet à toutes les évaluations externes sur la qualité de l'enseignement. Or, les établissements d'enseignement sont tenus de divulguer sur leur page d'accueil les moyennes générales de tous les examens finals. Les écoles publiques et les grammar schools ayant un programme d'enseignement national commun, la comparaison entre les moyennes générales de ces écoles est facile.

Le ministère de l'Éducation cherche au demeurant à encourager cette comparaison en autorisant, une fois par année, la consultation en ligne des moyennes générales des écoles publiques et privées. Même si le ministère se défend que l'établissement d'un palmarès fasse partie de sa politique (Andersen et Dahler-Larsen, à paraître), les journaux danois dressent des palmarès d'après les données publiées sur son site. Mais la présentation des données au nom des journaux et non du ministère enlève tout caractère officiel à ces palmarès (Andersen et Dahler-Larsen, à paraître).

Seules les moyennes générales peuvent servir de références à des fins de comparaison entre les établissements d'enseignement. L'information prospective et l'information évaluative ne sont pas disponibles à cet effet. Dans une large mesure, chaque établissement est libre de choisir l'information qu'il souhaite rendre publique et sous quelle forme. La loi danoise sur la transparence et l'ouverture en éducation leur accorde

cette marge de manœuvre sans pour autant que les établissements puissent contrôler et interpréter l'ensemble des données. L'équilibre entre l'information utile aux comparaisons et celle qui ne l'est pas dépend de la structure des données institutionnalisée par la loi qui permet une approche plus nuancée (Andersen et Dahler-Larsen, à paraître).

## ■ L'AUTOÉVALUATION

La loi sur la transparence et l'ouverture incite, et même parfois oblige, les établissements d'enseignement de tous niveaux à se doter de modes d'autoévaluation. En 2004, l'OCDE mentionnait dans son rapport sur l'éducation primaire au Danemark que le concept d'évaluation était certes intégré au système d'éducation danois, mais qu'il ne semblait pas avoir été implanté dans les pratiques des écoles danoises (The Ministry of Education, 2004, 65, 129). Cette observation laisse supposer que le Danemark accuse un retard en matière d'évaluation comparativement à d'autres pays (Hansen et Hansen, 2003).

Il est nécessaire de clarifier le concept d'autoévaluation en établissant une distinction entre évaluation interne et évaluation externe. Aucune définition acceptée de l'expression « évaluation interne » ne figure dans la documentation scientifique sur l'évaluation. Pour certains, distinguer entre évaluation interne et évaluation externe revient à répondre à la question suivante : qui devrait organiser, effectuer et utiliser l'évaluation ? (Vedung, 1997, 115). Pour d'autres, il s'agit de faire la distinction entre ceux et celles qui sont engagés dans l'évaluation (personnes internes et externes) et où se situe le contrôle du processus (au sein ou à l'extérieur de l'organisme fournisseur de services) (Qureshi, 1998, 137). Des personnes internes et externes interviennent sans aucun doute dans l'évaluation. Les évaluations externes peuvent être organisées et utilisées par des personnes internes, mais elles doivent, par définition, être effectuées par des personnes externes (Vedung, 1997, 116).

Selon Vedung (1997, 116), une évaluation interne est une autoévaluation. Quant à Love (1998, 145), il définit l'évaluation interne comme une forme spécifique d'évaluation qui est partie intégrante du modèle et du développement organisationnel et dans laquelle les membres du personnel agissent comme évaluateurs

contrairement aux évaluations externes axées sur le contrôle (Love 1998, 145). Dans ce sens, l'autoévaluation est associée à une évaluation axée sur l'apprentissage et le développement et le groupe cible est principalement un groupe interne de l'organisation. De ce point de vue, l'autoévaluation a une connotation positive et reflète la démocratie et la réceptivité. Toutefois, il peut s'avérer plus fructueux de ne pas inclure de considérations normatives dans la définition de l'autoévaluation et de mettre plutôt l'accent sur son organisation (Dahler-Larsen, 2003, 30, 32; Vedung, 1997).

En encourageant les établissements d'enseignement à «s'autoévaluer» le gouvernement s'est fixé pour objectif de généraliser une culture axée sur les résultats et l'évaluation dans le secteur danois de l'éducation. L'établissement d'enseignement garde la maîtrise de l'élaboration, de l'application et de l'utilisation de l'évaluation. Les établissements visés appartiennent à l'enseignement primaire, secondaire et universitaire. Ce sont des établissements d'enseignement autonomes qui reçoivent des subventions publiques.

Le concept d'autoévaluation apparaît dans plusieurs publications du ministère de l'Éducation du Danemark. Sa définition manque parfois de clarté et on la confond fréquemment avec l'évaluation interne. Que recouvre au juste l'autoévaluation et pourquoi les établissements d'enseignement devraient-ils l'utiliser sont des questions rarement abordées, si ce n'est dans des généralités sur l'accroissement de la qualité de l'éducation. L'autoévaluation est associée à une assurance qualité et elle fait le plus souvent référence à l'évaluation de l'enseignement (The Danish Government, 2002; The Danish Ministry of Education, 2003).

Sans être coercitive, l'approche danoise pour généraliser la pratique de l'autoévaluation dans le secteur de l'éducation peut être qualifiée d'insistante, ne serait-ce qu'à travers l'action de l'Institut danois d'évaluation. On peut cependant s'interroger sur la capacité des établissements d'enseignement à disposer des outils nécessaires sur le plan méthodologique et organisationnel pour réaliser une autoévaluation fondée sur une évaluation rétrospective et systématique de leurs résultats, évaluation destinée à jouer un rôle dans des situations pratiques futures (Vedung, 1997, 3). Par ailleurs, l'autoévaluation reste un outil à la disposition

des autorités nationales pour contrôler la performance des établissements d'enseignement et leur dicter, le cas échéant, des mesures autodisciplinaires.

## ■ RESPONSABILITÉ ET ÉVALUATION

Le système d'évaluation mis en place dans le secteur danois de l'éducation est en grande partie fondé uniquement sur des indicateurs. Il se présente comme un système ouvert qui laisse une place à l'interprétation sur certains points : en quoi l'évaluation est-elle importante ? Qui effectue une évaluation ? Comment ? La démarche qualité «via l'évaluation» n'est pas spécifique au secteur de l'éducation du Danemark. De nombreux pays occidentaux ont mené dans leur propre secteur de l'éducation des réformes fondées sur l'évaluation et la performance (Cibulka et Derlin, 1996; Ladd, 1996; Newman, King et Rigdon, 1997; Schwartz, 2000 et 2002). Toutefois, les systèmes d'évaluation diffèrent d'un pays à l'autre en raison des diverses définitions de la performance en éducation et des solutions retenues qui respectent les singularités nationales.

À titre d'exemple, l'introduction de l'évaluation dans un système de responsabilité auquel se greffe un système distinct de causalité et des mesures incitatives vise à s'assurer que les établissements d'enseignement sont responsables de la prestation des services, de ses progrès et de l'accroissement de sa performance. «La responsabilité repose sur la croyance voulant qu'une fois les objectifs clairement définis et formulés en des termes comportementaux, ils peuvent être mesurés et que le financement peut être accordé en fonction de ces résultats» (Cibulka et Derlin, 1996, 481). Ici, l'évaluation renforce la responsabilité externe des établissements publics et des fonctionnaires (Ladd, 1996; Schwartz, 2000, 14).

Avec la publication de l'information évaluative et des résultats de la performance des établissements d'enseignement, les pouvoirs publics présupposent que les établissements apporteront des correctifs à leur action conformes aux recommandations formulées dans l'évaluation. Ils espèrent également que les établissements se rapprocheront des normes nationales d'une part et surtout des attentes de leur clientèle d'autre part, celle-ci pouvant se révéler volatile et choisir un établissement concurrent. Cette auto-discipline ou autoréglementation est une obligation

pour les institutions publiques (Scott, 2004, 164). Dans certains cas, comme en Finlande, le financement ou même la rémunération des enseignants dépendent de la performance de l'établissement.

Le modèle de causalité repose sur l'hypothèse voulant que les employés des établissements d'enseignement atteignent plus efficacement les objectifs de performance lorsque ces objectifs sont clairs, lorsque l'information sur le degré de réussite est accessible et lorsque des mesures incitatives pour atteindre les objectifs sont réellement appliquées (Newman, King et Rigdon, 1997, 43). L'examen public de la performance est dans ce cas un gage d'efficacité et de prise en compte des besoins des clients. Ce modèle reste encore un idéal à atteindre car les objectifs cibles, le taux de réussite et les mesures incitatives sont presque toujours le résultat de compromis politiques et manquent donc de transparence (Andersen, à paraître).

Bien que le système danois d'évaluation possède des caractéristiques similaires aux systèmes d'évaluation de pays tels que les États-Unis et le Royaume-Uni, la notion de responsabilité est, au Danemark, absente du discours politique sur la performance scolaire. Sans véritable équivalent dans le vocabulaire danois, le terme « responsabilité » chapeaute de vagues notions de transparence, d'ouverture, de qualité et de compétition.

Le seul élément du système danois d'évaluation comparable aux mesures de la performance et à la divulgation de l'information aux États-Unis et au Royaume-Uni est en fait la publication des moyennes générales des examens finals telle qu'elle est prévue dans la loi sur la transparence et l'ouverture. En revanche, le système danois fait peu de place aux mesures incitatives ce qui confère une ambiguïté aux taux de réussite présentés. C'est pourquoi la responsabilité des établissements est engagée au sujet de l'information parue sur Internet.

En résumé, le système danois d'évaluation se concentre non seulement sur la performance ou les résultats en éducation, mais aussi sur la mise en place de processus organisationnels plus complexes tels que l'apprentissage organisationnel. Bien qu'il épouse une tendance internationale, il diffère considérablement des systèmes d'évaluation des autres pays. Tous les systèmes partagent toutefois une caractéristique commune : les recommandations et les résultats de

l'évaluation sont utilisés.

Le concept du paradoxe de la performance doit attirer notre attention sur le fait que l'évaluation peut entraîner des conséquences imprévues, par exemple invalider la performance ou avoir une influence négative sur elle (Van Thiel et Leeuw, 2002). Ce sont les effets constitutifs de l'évaluation qui génèrent ces paradoxes. Ces effets valorisent l'aspect « utilitariste » de l'évaluation par rapport à l'utilisation traditionnelle d'une information. Les effets constitutifs mettent en évidence les nouvelles réalités apparues à la suite des réactions en chaîne provoquées par la publication de l'information évaluative et des résultats, tant à l'intérieur qu'à l'extérieur de l'établissement d'enseignement (Andersen et Dahler-Larsen, à paraître).

La question reste en suspens : les établissements d'enseignement et leurs employés vont-ils modifier leur comportement après l'instauration du nouveau système d'évaluation au Danemark ? À cause de la publication d'information évaluative, les enseignants devront peut-être adapter directement ou indirectement leur enseignement en visant avant tout à obtenir pour leurs élèves de bons résultats aux examens finals ou aux tests nationaux (Schwartz, 2002). La publication de ces évaluations favorisera-t-elle les processus internes de développement de la qualité ? Le lien n'est pas établi pour l'instant (Andersen et Dahler-Larsen, à paraître).

En choisissant stratégiquement l'information à diffuser sur leur page web, les établissements d'enseignement pourraient par ailleurs être tentés de peaufiner leur image (Andersen et Dahler-Larsen, à paraître). Une diminution de la collaboration entre les établissements et de leur partage des bonnes pratiques peut enfin se révéler un effet secondaire non désiré de la priorité accordée à la publication d'information évaluative.

## ■ CONCLUSION

On a procédé au sein du secteur danois de l'éducation à une formalisation de l'évaluation de la qualité de l'éducation. Les autorités nationales ont choisi de superviser leur système d'éducation décentralisé à travers l'évaluation de la performance. Les procédures mises en place s'appliquent pour l'essentiel aux niveaux inférieurs. Les gouvernements tentent par ailleurs d'établir des processus organisationnels plus complexes pour encourager les établissements

d'enseignement à instaurer une culture locale de la performance et de son évaluation.

Les systèmes d'évaluation retenus doivent tenir compte de l'autonomie traditionnellement accordée aux établissements et aux structures de l'éducation danoise. Ils tendent néanmoins à rétablir l'équilibre entre l'échelon central et les échelons décentralisés ainsi qu'entre des modèles d'évaluation normalisés et des modèles plus autonomes.

L'Institut danois d'évaluation et la loi sur la transparence et l'ouverture contribuent à une certaine forme de centralisation et de standardisation de l'évaluation tout en permettant aux établissements d'enseignement de conserver une part d'autonomie. En particulier, le fait d'encourager l'autoévaluation accentue cette autonomie. Avec la généralisation de l'évaluation au niveau local et une harmonisation au niveau national, le gouvernement danois entend relever le défi de la qualité de l'éducation et obtenir pour le pays de meilleurs résultats dans les comparaisons internationales.

## BIBLIOGRAPHIE

- The Act on transparency and openness in education*. Loi n° 414 du 6 juin 2002.
- The Act on the Danish Evaluation Institute*. Loi n° 775 du 24 juin 2005.
- Andersen, V. N. (à paraître). « Transparency and openness in Education », *Scandinavian Political Studies*.
- Andersen, V. N. et P. Dahler-Larsen (à paraître). « The Framing of Public Evaluation Data. Transparency and Openness in Danish Schools », dans J. D. Breul, R. Boyle et P. Dahler-Larsen (eds.). *Bringing Evaluation into Public Life*, New Brunswick, Transaction Publishers.
- Beck, U., A. Giddens et S. Lash (1994). *Reflexive Modernization. Politics, tradition and aesthetics in the modern social order*. Cambridge, Polity Press.
- Christensen, J. G. (2000). « Governance and devolution in the Danish school system », p. 198-216 dans M. A. Arnott et C. D. Raab (eds.). *The Governance of Schooling. Comparative studies of devolved management*, London, Routledge.
- Cibulka, J. G. et R. L. Derlin (1996). « State educational performance reporting policies in the U.S.: Accountability's many faces », *International Journal of Education Research*, 23, p. 479-492.
- Dahler-Larsen, P. (2003). « Hvad er Selvevaluering? » dans P. Dahler-Larsen (red.), *Selvevalueringens hvide sejl*, Odense, Syddansk Universitetsforlag.
- Dahler-Larsen, P. (1998). *Den Rituelle Refleksion*, Odense, Syddansk Universitetsforlag.
- The Danish Ministry of Education (2003). *De gymnasiale uddannelser. Redegørelse til Folketinget*, Copenhagen, The Danish Ministry of Education.
- The Danish Government (2002). *Better Education. Action plan*, Copenhagen, The Danish Ministry of Education.
- Hansen, H. F. et A. Hansen (2001). Evaluering i Danmark p. 155-176 dans Carsten Greve (ed.), *Offentlig Ledelse og styring i bevægelse*, København, Jurist- og Økonomforbundets Forlag.
- Jordana, J. et D. Levi-Faur (2004). « The Politics of Regulation in the Age of Governance » p. 1-28 dans J. Jordana et D. Levi-Faur (eds.), *The Politics of Regulation. Institutions and Regulatory Reforms for the Age of Governance*, Cheltenham, Edward Elgar.
- Knudsen, T. (2001). « Undervisningsministeriet – Fra klientministerium til New Public Management » p. 33-47 dans J. Glerup (ed.), *Voksenuddannelse under forandring*, København, Gads Forlag.
- Kristoffersen, D. (2003). « Denmark » p. 25-34 dans The Danish Evaluation Institute (ed.), *Educational Evaluation around the World: An International Anthology*, Copenhagen, The Danish Evaluation Institute.
- Ladd, H. F. (1996). « Introduction » dans H. F. Ladd (ed.), *Holding Schools Accountable. Performance-Based Reforms in Education*, Washington D.C., The Brookings Institution.
- Leeuw, F. L. (2003). « Trends, Topics and Theories » p. 9-22 dans The Danish Evaluation Institute (ed.), *Educational Evaluation around the World: An International Anthology*, Copenhagen, The Danish Evaluation Institute.

- Love, A. J. (1998). Internal Evaluation: Integrating Evaluation and Social Work Practice. *Scandinavian Journal of Social Welfare*. Vol. 7, no 2, avril 1998, pp. 145-152.
- Newman, F., B. M. King et M. Rigdon (1997). «Accountability and School Performance: Implications from Restructuring Schools», *Harvard Educational Review*, vol. 67, n° 1, p. 41-69.
- Norrild, P. (2004). «Fra modstand til selvansagelse», *Uddannelse*, vol. 37, n° 6, p. 16-23.
- OECD (2005). *Education at a Glance. OECD Indicators 2005*, Paris, OECD.
- Pierre, J. et B. G. Peters (2000). *Governance, Politics and the State*, Houndmills, Macmillan.
- Qureshi, H. (1998). «Internal and External evaluation of social work», *Scandinavian Journal of Social Welfare*, vol. 7, p. 137-144.
- Royal Danish Ministry of Foreign Affairs (2003). *Factsheet Denmark. Education and Training*. Copenhagen, The Royal Danish Ministry of Foreign Affairs.
- Scott, C. (2004). «Regulation in the age of governance: the rise of the post-regulatory state» p. 145-174 dans J. Jordana et D. Levi-Faur (eds.), *The Politics of Regulation. Institutions and Regulatory Reforms for the Age of Governance*, Cheltenham, Edward Elgar.
- Schwartz, R. (2002). «New Public Management Education Reform: Politics, Policy Image and Political-Administrative Systems», *International Public Management Journal*, vol. 5, n° 3, p. 277-295.
- Schwartz, R. (2000). «School Accountability – An Elusive Policy Solution. The Israeli Experience in Comparative Perspective», *Journal of Public Policy*, vol. 20, n° 2, p. 195-218.
- Van Thiel, S. et F. L. Leeuw (2002). «The performance paradox in the public sector», *Public Performance and Management Review*, vol. 25, n° 3, p. 267-281.
- Vedung, E. (1997). *Public Policy and Program Evaluation*, New Brunswick, Transaction Publishers.



la première source d'information  
sur l'État québécois,  
son organisation,  
son fonctionnement,  
son action



**RASSEMBLE**  
.....

les données les plus récentes  
et l'information disponible  
depuis les 25 dernières années  
sur la gouvernance publique  
au Québec



**DONNE ACCÈS**  
.....

gratuitement à une information  
exhaustive sur l'État québécois  
(organisation politico-administrative,  
effectif du secteur public et finances  
publiques), les politiques publiques  
et la gestion publique



**PRÉSENTE ET  
ANALYSE**  
.....

les données quantitatives et  
qualitatives recueillies auprès  
des institutions gouvernementales  
et des universités

**COMPARE**  
.....

l'organisation, le fonctionnement  
et l'action de l'État québécois  
avec ceux d'autres entités étatiques,  
souveraines ou non

**RÉACTUALISE**  
.....

l'information en continu  
par la mise en place d'un service  
de veille spécifique



**FACILITE**  
.....

la réflexion à conduire  
sur l'action publique

Un site pour les spécialistes de l'administration  
publique (praticiens, chercheurs, étudiants)  
et tous les citoyens

# Procédures et systèmes gouvernementaux d'évaluation de la performance publique: les leçons australiennes

Par **Keith Mackay**<sup>1</sup>, agent principal d'évaluation au Département de l'évaluation des opérations de la Banque mondiale, coordonnateur du développement de la capacité d'évaluation • kmckay@worldbank.org

Traduit de l'anglais

De plus en plus, il est reconnu qu'une orientation sur les résultats est un élément important d'une bonne gouvernance et qu'il faut insister sur les résultats pour atteindre des niveaux élevés de performance. Un certain nombre de pays optent pour cette orientation, et ce, en établissant ou en bonifiant les fonctions de suivi et d'évaluation (S&É) de leurs activités gouvernementales. Le présent article décrit en les rapprochant deux générations de systèmes d'évaluation de la performance et de la gestion appliqués en Australie au cours des 18 dernières années. L'article offre aussi des enseignements en vue de l'édification de tels systèmes dans d'autres pays, notamment les avantages et les inconvénients d'une approche centralisée par rapport à une approche plus décentralisée.

## ■ LA PREMIÈRE GÉNÉRATION: DE 1987 À 1997

### Priorité à la réforme du secteur public

Le gouvernement travailliste élu en 1983 avait l'intention de mettre en œuvre une série de réformes du secteur public en vue d'améliorer sa performance de façon significative. Ces réformes consistaient, entre autres, à «laisser les gestionnaires gérer» grâce à une dévolution en leur faveur des pouvoirs et des responsabilités. À travers cette plus

grande autonomie des gestionnaires, on visait une meilleure performance. En même temps, le gouvernement militait pour des principes de gestion de programmes et de budgétisation axés sur l'efficacité et l'efficience au moyen d'un assainissement des pratiques, de la collecte d'information sur la performance et de la généralisation de l'évaluation de programmes.

Le ministère des Finances fut l'architecte principal d'un bon nombre des réformes du secteur public, ce qui a contribué à renforcer son rôle de coordonnateur du budget et de superviseur des dépenses. Le ministère des Finances cherchait par la même occasion à s'affranchir des questions «domestiques» de dépenses pour se consacrer aux questions de politique financière de niveau supérieur, comme le montrent les analyses de sa politique et les exposés préparés à l'appui du processus budgétaire annuel. En matière de dépenses budgétaires, il s'inquiétait non seulement de la nécessaire priorité de réduire les dépenses gouvernementales, mais aussi de la recherche de moyens pour que les dépenses soient plus efficaces et plus efficaces.

En 1988, un examen mené par le ministère des Finances sur les pratiques d'évaluation des ministères avait permis de déceler:

- un manque d'intégration de l'évaluation dans la prise de décisions ministérielles et financières;
- une tendance pour les évaluations à se concentrer sur les questions d'efficacité et de procédure au

<sup>1</sup> Keith Mackay est agent principal d'évaluation au Département de l'évaluation des opérations de la Banque mondiale. Il est aussi le coordonnateur du développement de la capacité d'évaluation. Il était auparavant le gestionnaire du système national de S&É du gouvernement australien. Les points de vue exprimés dans le présent article n'engagent que l'auteur et ne représentent pas nécessairement le point de vue de la Banque mondiale.



lieu d'aborder la question plus fondamentale de l'efficacité globale des programmes, à savoir l'atteinte effective de leurs objectifs;

- un faible niveau de compétences en évaluation et de capacité analytique;
- un manque de clarté sur le rôle des ministères centraux dans l'évaluation, en particulier du ministère des Finances.

### La stratégie d'évaluation du gouvernement

Le ministère des Finances en a conclu qu'il était insuffisant de «laisser les gestionnaires gérer». Il a été jugé plus judicieux de «faire en sorte que les gestionnaires gèrent». Par conséquent, vers la fin de l'année 1988, le ministre des Finances a obtenu l'accord du Cabinet pour élaborer une stratégie d'évaluation formelle ayant pour principe sous-jacent la responsabilité des ministères dans la détermination des priorités d'évaluation, la préparation des plans d'évaluation et la conduite des évaluations.

La stratégie visait trois objectifs: fournir tout d'abord de l'information fondamentale sur la performance du programme afin d'aider la prise de décisions du Cabinet et l'établissement de ses priorités, en particulier dans le processus budgétaire annuel lorsqu'un grand nombre de propositions concurrentes sont mises de l'avant par les ministères; encourager ensuite les gestionnaires de programmes au sein des ministères à utiliser l'évaluation pour améliorer la performance de leurs programmes; renforcer, enfin, la reddition de comptes dans un cadre décentralisé en offrant une preuve formelle de la maîtrise et de la gestion des ressources des programmes par les gestionnaires de ces programmes.

La stratégie d'évaluation était rendue opérationnelle par quatre grandes exigences pour les ministères:

- que chaque programme soit évalué tous les trois à cinq ans;
- que chaque ministère prépare un plan d'évaluation ministériel annuel (PEP) avec une couverture prospective sur trois ans et le soumette au ministère des Finances.

· que les nouvelles propositions de politiques des ministères incluent une déclaration des arrangements proposés pour l'évaluation future de ces politiques;

· que les rapports d'évaluation terminés soient normalement publiés, sauf en cas de considérations importantes à cause du caractère délicat des politiques ou de la sécurité nationale et que les conclusions d'évaluation soient incluses dans le document budgétaire des ministères destiné au Parlement.

Le Cabinet a aussi convenu que le ministère des Finances aurait l'occasion de faire ses commentaires sur les PEP et sur le mandat des évaluations pour garantir leur cohérence avec les politiques et les priorités à l'échelle gouvernementale. Le ministère des Finances serait prêt à participer directement à certaines évaluations, sous réserve de négociations entre lui et le ministère sectoriel.

Au cours des années suivantes, des rapports d'un comité parlementaire et du Bureau national de vérification ont noté un manque d'uniformité constant dans la couverture et la portée de l'activité d'évaluation dans les ministères. Les deux rapports ont soutenu que le ministère des Finances devrait adopter une approche plus active pour encourager les ministères à planifier et à entreprendre l'évaluation. Le Ministère a ensuite créé une direction distincte chargée d'offrir des conseils, du soutien, de la formation et de l'encouragement en matière d'évaluation aux autres ministères et aussi au sein de ses propres services.

L'examen continu par le ministère des Finances de la stratégie d'évaluation a révélé une lacune: on ne prêtait pas suffisamment attention à la collecte, à l'utilisation et à la déclaration de façon régulière de l'information sur la performance, par le biais d'outils comme les systèmes d'information de gestion, les indicateurs de performance et la fixation de cibles de performance (Mackay, 1998). Cette situation peut sembler quelque peu paradoxale parce que l'évaluation fait intervenir des techniques relativement sophistiquées. À ce titre, elle était en général menée correctement. Toutefois, la fixation d'objectifs de programme et la collecte d'information de performance sur une base régulière sont souvent plus

faciles à conceptualiser qu'à réaliser. Une différence tenait au fait que l'évaluation avait été ordonnée, alors que la collecte d'informations sur la performance ne l'avait pas été.

Ainsi, en 1995, le ministère des Finances a obtenu l'accord du Cabinet pour préparer une série courante d'examens complets, étalés sur trois ans, des objectifs de programmes et de l'information sur la performance de tous les programmes dans tous les ministères. Ces examens ont été effectués de façon conjointe par le ministère des Finances et par chaque ministère sectoriel, les résultats étant communiqués aux ministres respectifs et au Cabinet plénier. Les examens ont marqué le début de la priorité accordée à l'information de performance qu'on retrouve dans la deuxième génération de réformes qui ont commencé en 1997.

### **Quel a été le succès de la stratégie d'évaluation ?**

• **La planification de l'évaluation.** Depuis le lancement de la stratégie, tous les ministères avaient préparé des PEP visant à promouvoir les évaluations au sein du ministère lui-même et dans ses organismes externes. L'exigence formelle, exprimée par le Cabinet, en vue de la création de PEP a certainement incité les ministères à préparer sérieusement des plans d'évaluation. Dès le milieu des années 1990, quelques 160 évaluations étaient en cours. La majorité d'entre-elles apparaissait importante du fait que les programmes comportaient des incidences majeures sur le plan de la politique et des dépenses. Les ministères décidaient eux-mêmes quels programmes devaient être inclus dans leur PEP et aussi quelles questions devaient être couvertes par le mandat d'évaluation. Toutefois, le ministère des Finances cherchait à influencer sur les choix des priorités d'évaluation des ministères. Quand il faisait ses suggestions, le Ministère cherchait autant à prévoir qu'à faire émerger les besoins d'information du Cabinet. Le ministère des Finances approchait parfois directement le Cabinet pour obtenir son aval au sujet des thèmes d'évaluation proposés et aussi du détail des mandats.

Dans sept vérifications d'efficacité et dans deux guides sur les « meilleures pratiques » en matière d'éva-

luation de programmes et d'information sur la performance, l'Australian National Audit Office (ANAO) a insisté auprès des ministères sur l'importance d'une planification systématique de leurs activités d'évaluation (ANAO 1993, 1997). De fait, l'appui solide de l'ANAO en faveur de l'évaluation, qui perdure encore aujourd'hui, et son programme dynamique de vérification de la performance – ce qui constitue en soi une forme d'évaluation – ont fourni un soutien précieux aux deux générations de systèmes d'évaluation de la performance de l'Australie.

L'obligation voulant que tous les programmes soient évalués tous les trois à cinq ans a également contribué à créer un contexte dans lequel l'évaluation serait la norme et non l'exception. L'idée de « balayage » périodique et complet des programmes impliquait une approche planifiée et échelonnée en matière d'évaluation. Cette exigence ne devait toutefois pas être acceptée comme telle. Il est très rare en effet que la totalité d'un programme soient compris dans une même évaluation. À l'inverse, il est courant qu'une évaluation se concentre sur certains problèmes ou aspects clefs d'un programme. Le défi consiste donc à s'assurer que ces questions difficiles soient effectivement évaluées. C'est le rôle joué par le ministère des Finances de façon active par le biais de la persuasion et de la participation directe aux évaluations.

• **La conduite de l'évaluation.** La plupart des ministères ont choisi de former des unités d'évaluation pour coordonner leur planification de l'évaluation formelle. Ces unités comprenaient au minimum deux ou trois personnes. Dans certains ministères, comme celui de l'emploi, il existait une direction distincte de 20 à 25 personnes chargées de la planification de l'évaluation, de l'offre de conseils sur la méthodologie d'évaluation, de la participation au comité directeur et de l'organisation d'un certain nombre de grandes évaluations, en particulier dans le domaine des programmes du marché du travail.

Il n'y avait pas d'approche normalisée dans les ministères en ce qui concerne la conduite des évaluations. Certaines évaluations réunissaient une vaste gamme d'intervenants externes et internes, que ce soit comme participants à un comité directeur d'évaluation ou, plus rarement, à l'équipe d'évaluation. D'autres étaient menées par une unité

d'évaluation centrale, mais il était plus habituel que la responsabilité des évaluations incombe au secteur en charge du programme évalué. Pour les évaluations importantes, à savoir celles énumérées dans les plans d'évaluation ministériels, il était fait appel fréquemment à un intervenant externe. Pour les programmes ayant des incidences majeures sur le plan des ressources ou des politiques, le ministère des Finances imposait le plus souvent sa présence dans le processus.

Une enquête du Bureau national de vérification a permis de conclure que sur l'ensemble des évaluations menées entre 1995 et 1997, environ la moitié portaient sur la prestation de services et la livraison de produits à des clients externes, tandis que 30 % étaient liées à des questions internes au ministère. Un tiers des évaluations portaient sur le caractère approprié des programmes, nouveaux ou établis, tandis que 15 % visaient l'élaboration de conseils pour le gouvernement (ANAO 1997).

Le grand nombre d'évaluations en cours en tout temps et le fait que plus de 530 rapports d'évaluation ont été publiés entre 1993 et 1997 témoignent de l'existence d'une grande activité d'évaluation au gouvernement australien, ce qui a permis de créer une « bibliothèque » de plus en plus fournie sur les conclusions d'évaluation.

• **La qualité de l'évaluation.** La qualité des rapports d'évaluation constitue une dimension plus difficile à mesurer. La rigueur des évaluations de programmes dépend de l'expertise et de l'objectivité des évaluateurs. L'ANAO a jugé la qualité d'un échantillon de rapports d'évaluation en 1997 et en a conclu que plus d'un tiers d'entre eux souffraient de faiblesses méthodologiques. Il est vrai que quelques évaluations publiées étaient de piètre qualité et on soupçonne que certaines d'entre elles avaient été produites à des fins intéressées, de façon à justifier le maintien ou l'extension d'un programme.

La perception du ministère des Finances était que la qualité des évaluations variait sensiblement, ce qui ne manquait pas de poser des problèmes quand le destinataire de l'évaluation était le Cabinet (aux fins de prise de décisions) ou le Parlement (aux fins de reddition de comptes). Dans ces cas-là, le Ministère informait le destinataire qu'il estimait que l'évaluation n'était pas fiable. Les ministères cherchaient en

général à éviter une telle critique pour s'épargner les remontrances du Cabinet.

Le Bureau national de vérification a toujours soutenu l'idée d'instituer des procédures centralisées de surveillance pour garantir la qualité des évaluations menées par les secteurs au sein d'un ministère. Quelques ministères ont retenu cette approche et elle a prouvé son efficacité. Mais la majorité des ministères ont préféré s'appuyer sur la connaissance que détenait le personnel sur le programme évalué et sur son engagement par appropriation des conclusions de l'évaluation. En la matière, le ministère des Finances avait comme stratégie de chercher à obtenir les avantages de l'autoévaluation tout en garantissant, par sa participation aux comités directeurs des évaluations importantes, une objectivité et une rigueur indiscutables.

Un inconvénient de l'approche décentralisée tenait au manque de compétences en évaluation et au manque d'expérience dans la tenue d'évaluations des gestionnaires de programme et de leur personnel. Une formation de base était cependant dispensée par le ministère des Finances. Du matériel pédagogique était distribué dans les services. On pouvait aussi faire appel à la communauté des consultants en évaluation de Canberra qui comprenait nombre d'universitaires spécialistes d'un domaine (par exemple la santé) ou experts en matière de recherche et d'analyse. Malgré cela, l'étude de l'ANAO de 1997 a montré que 20 % des ministères s'inquiétaient du manque de formation en techniques d'évaluation avancées et il semble que, dans ce manque, résidait le maillon faible du système d'évaluation australien.

• **L'utilisation de l'évaluation.** L'utilisation des conclusions d'évaluation est le facteur décisif. Si elle est aléatoire ou incohérente, les évaluations sont inutiles. Il a été prouvé que les évaluations avaient servi de référence dans le cadre du processus budgétaire au moment de l'élaboration d'options de politiques et de leur examen par le Cabinet. Considérés comme des spécialistes, les fonctionnaires du ministère des Finances qui assistaient aux réunions du Cabinet portant sur des questions budgétaires ont été priés de témoigner sur l'étendue de l'influence des évaluations sur les propositions budgétaires des ministres.

Dans le budget 1990-1991, on estime que près de 230 millions de dollars australiens (près de 175 millions de dollars américains) de propositions nouvelles soumises par les ministres ont été directement ou indirectement influencées par les conclusions d'une évaluation. Dès 1994-1995, ce montant est passé à 2300 millions de dollars australiens, soit, en proportion des politiques proposées, un bond de 23 % à 77 %.

Ces résultats spectaculaires soulignent l'importance que les ministres et les fonctionnaires attachaient à l'existence de conclusions d'évaluation dans la préparation des nouvelles propositions de politiques. Il était très important d'avoir le soutien actif des principaux ministres du Cabinet. Cet appui a incité les ministères à planifier et à mener des évaluations de grande qualité. Il s'est également manifesté à travers les nombreuses demandes d'évaluation émanant du Cabinet lui-même.

Il a aussi été démontré que les évaluations peuvent avoir une influence non négligeable sur les arbitrages budgétaires (essentiellement les coupures de programmes) présentés par le ministère des Finances ou par les ministères sectoriels pour examen par le Cabinet. En 1994-1995, des arbitrages représentant 500 millions de dollars australiens (65 % du total) ont été significativement influencés par les conclusions d'évaluation.

La nature même du système budgétaire du gouvernement australien a permis cette avancée de l'évaluation. L'Australie dispose en effet d'un mécanisme d'élaboration de politiques qui assure la transparence des coûts des politiques concurrentes et encourage le débat et la consultation entre les intervenants au sein du gouvernement. Dans ce « marché des idées », les conclusions d'évaluation apparaissaient parfois comme un avantage pour certains face à leurs contradicteurs.

Bien entendu, il faut être conscient des limites de l'influence des évaluations sur la prise de décisions des ministres ou du Cabinet. Les conclusions d'évaluation peuvent révéler l'efficacité d'un programme

dans la poursuite de ses objectifs, mais il est rare qu'elles permettent de conclure définitivement sur l'utilité, ou l'inutilité, globale de l'activité. Cette décision conserve son caractère politique.

Les fonctionnaires du ministère des Finances ont par ailleurs été sondés sur l'influence de l'évaluation sur les décisions du Cabinet dans les budgets 1993-1994 et 1994-1995. Le rôle de l'évaluation a été jugé essentiel. En 1994-1995 par exemple, on estime que 68 % des 3 740 millions de dollars australiens de propositions étudiées (les nouvelles politiques plus les propositions d'économie) résulte de cette influence.

Campbell (2001) souligne que le ministère des Finances était un point névralgique dans la politique conduite par le Cabinet en matière d'évaluation. Ce ministère avait joué un rôle de premier plan dans la réforme du secteur public tout au long des années 1980. Ses initiatives portaient sur la dévolution du pouvoir en matière de gestion des coûts aux ministères et sur la promotion d'une culture d'évaluation qui permettaient de suivre et de juger l'efficacité avec laquelle les ministères utilisaient ce pouvoir discrétionnaire de dépenser.

Le vérificateur général a ajouté qu'à son avis [traduction libre] « le succès de l'évaluation au palier fédéral de gouvernement... a été largement dû à sa pleine intégration dans les processus budgétaires. Là où il y avait un engagement de ressources, une certaine forme d'évaluation était nécessaire pour justifier pratiquement l'ensemble des soumissions budgétaires » (Barrett, 2001). En comparant le succès relatif des approches de l'Australie et du Canada concernant l'incorporation de l'évaluation dans le processus budgétaire, Schick (2001) remarque :

**« Le Canada a déployé de vastes efforts autour du Contrôleur général dans les années 1970 et l'Australie a adopté, à la fin des années 1980, une stratégie d'évaluation ambitieuse. Les efforts du Canada n'ont guère été payants mais l'Australie a procédé à une réaffectation importante de ses ressources budgétaires. On pense que le Canada a**

<sup>2</sup> Le Canada a introduit un nouveau cadre de gestion des résultats dans l'ensemble du gouvernement en 2000, ce qui inclut une stratégie renouvelée d'évaluation avec un financement supplémentaire de 20 millions de dollars canadiens sur quatre ans pour aider les ministères à renforcer leurs capacités en matière d'évaluation. Voir, par exemple, le Secrétariat du Conseil du Trésor du Canada (2000).

échoué parce qu'il a centralisé l'évaluation, affaiblissant ainsi la coopération des ministères dépensiers qui aurait pu être affectée négativement par les conclusions. On pense que l'Australie a réussi parce qu'elle a su impliquer largement les ministères concernés dans la conception et l'utilisation des évaluations.<sup>2</sup>»

L'enquête de l'ANAO de 1997 a enfin permis de conclure que l'incidence ou l'utilisation des évaluations par les ministères sectoriels a surtout été profitable en terme de gains d'efficacité opérationnelle et, dans une moindre mesure, pour la prise de décisions d'affectation de ressources et l'amélioration du service à la clientèle.

## ■ LA DEUXIÈME GÉNÉRATION : DE 1997 À NOS JOURS

### Le contexte

L'élection d'un gouvernement conservateur en mars 1996 a marqué un changement dans l'orientation de la réforme du secteur public en Australie. Le nouveau gouvernement avait une nette préférence idéologique en faveur du secteur privé qu'il trouvait plus efficace que le secteur public et il militait fortement pour une gestion axée sur les résultats. Il avait dit à quel point il jugeait la fonction publique bureaucratique et prisonnière de la « paperasserie », ce qui a par la suite conduit à une réduction importante du nombre de fonctionnaires<sup>3</sup>. Pour illustrer ce virage vers le secteur privé, les hauts dirigeants des ministères sont devenus des présidents directeurs généraux (PDG), responsables des résultats et non des processus bureaucratiques. Dans ce contexte, on jugeait inconcevable de limiter les actions des PDG par des contrôles administratifs excessifs, ce qui signifiait qu'il fallait « laisser les gestionnaires gérer », tel que le proclamait le principe adopté au début des années 1980.

Cet effort de simplification toucha également la préparation des plans ministériels d'évaluation. De 20 à

30 pages au départ, certains atteignaient alors 120 pages, avec une hausse concomitante de la charge de travail administrative pour les préparer. On comprenait bien que malgré l'importance des conclusions d'évaluation pour la prise de décisions par les gestionnaires de programmes et le Cabinet, il n'était pas nécessaire d'avoir des plans aussi détaillés et élégamment rédigés. En même temps, les ministères sectoriels réclamaient une diminution du contrôle du ministère des Finances et un allègement de la reddition de comptes à celui-ci.

### La nature des réformes

Ainsi, un nouveau système de gestion de la performance fut créé en 1997. Il combinait des principes, des attentes et des exigences formelles. Globalement, le nombre et la nature des exigences formelles étaient considérablement réduits ou simplifiés par rapport à ceux du gouvernement antérieur (Russell, 2003). Tous les ministères devaient mesurer leur performance en termes de résultats, en fonction d'indicateurs de performance relativement simples. Les ministères devaient fournir tous les ans au Parlement :

- une reddition de comptes *ex ante* sous forme de déclarations budgétaires ministérielles où le lien entre les affectations de dépenses prévues et la performance escomptée soit mis en évidence ;
- une reddition de comptes *ex post* sous forme de rapports annuels ministériels portant sur les dépenses réelles et les résultats réels.

La comptabilité d'exercice a été introduite au gouvernement fédéral dans son budget de 1999-2000. L'objectif était de garantir que les affectations budgétaires révèlent le vrai coût des activités gouvernementales. Il s'agissait aussi de faciliter l'étalonnage entre le coût unitaire des extrants et les résultats gouvernementaux dans la prestation de services.

Les montants en dollars des affectations budgé-

<sup>3</sup> Dans ce contexte, il convient de citer les commentaires d'Allen Schick (2003) : « L'éthique du service public constitue les fondations de la performance gouvernementale qui dépend au moins autant des individus que de la machine et des processus administratifs... Dans la plupart des pays développés, l'éthique du service public existe toujours, mais elle n'est pas aussi forte que par le passé. Elle a été affaiblie par des forces socio-économiques puissantes telles que... le recours accru aux marchés et aux sous-traitants privés pour assurer des services publics et la baisse de la considération dont jouissent les fonctionnaires. »

taires sont désormais fixés uniquement en fonction des résultats visés, ce qui contraste avec l'insistance traditionnelle sur le montant des dépenses (les intrants) ou sur celui des extrants (d'après l'approche de pays comme la Nouvelle-Zélande) et ce qui fait de l'Australie un cas singulier. Les PDG se sont vus octroyer le pouvoir de procéder à ces affectations. Leur obligation de rendre compte passe par une entente formelle de performance avec leur ministre. Les ministères et organismes qui fournissent des services directement au public sont tenus d'élaborer des chartes de service public indiquant des niveaux réels de performance, ceux-ci étant repris dans leurs rapports au Parlement.

Parallèlement, les exigences formelles d'évaluation des systèmes de la première génération ont été abandonnées, tout comme il a été mis un terme au soutien que le ministère des Finances, rebaptisé entre temps ministère des Finances et de l'Administration, avait accordé aux ministères sectoriels. Le « nouveau » ministère encourage encore les ministères sectoriels à évaluer les politiques et les programmes clefs, au mieux sur un cycle de cinq ans, mais il n'y a plus l'exigence formelle d'entreprendre des évaluations. Les soumissions de financement au Cabinet doivent être accompagnées d'une preuve qu'une évaluation a bien eu lieu. Rien de plus. Les conclusions de l'évaluation ne font plus systématiquement l'objet d'un rapport sauf dans de rares cas.

Toutefois, à compter de 2004-2005, les ministères qui souhaitent reconduire le financement (par une soumission au Cabinet) de programmes à durée fixe sont tenus d'examiner la performance de ces programmes pour mesurer l'atteinte des objectifs. Le ministère des Finances et de l'Administration a préparé des mandats génériques pour encadrer ces examens. Malheureusement, la même exigence n'existe pas pour les programmes permanents.

Le ministère des Finances et de l'Administration a caractérisé de « décentralisée » cette nouvelle génération du système d'évaluation de la performance et de la gestion (Russell, 2003). Le vérificateur général de l'Australie a parlé lui de « déréglementation de l'évaluation » (Barrett, 2001).

Hormis l'introduction de la comptabilité d'exercice que le gouvernement précédent avait élaborée durant son mandat, la principale innovation du nouveau gouvernement a consisté dans la création d'un « cadre des résultats et des extrants ». À compter du budget 1999-2000, le cadre exige des ministères qu'ils accordent leur prévision de résultats et l'engagement des extrants jugés nécessaires à l'obtention de ces résultats avec les objectifs gouvernementaux. Ce cadre prévoit aussi des mesures claires de la performance des ministères. Cette évaluation doit être réalisée par une collecte, une analyse et une déclaration régulières de l'information sur la performance.

Reflets de la stratégie d'évaluation antérieure, les trois objectifs du « cadre des résultats et des extrants » sont les suivants :

- épauler l'élaboration de la politique du gouvernement ;
- consolider la gestion interne des ministères, y compris la formation du personnel ;
- étayer la déclaration externe aux fins de reddition de comptes.

Sur le papier du moins, le « cadre des résultats et des extrants » est une initiative louable à bien des égards. En plus de ses trois objectifs, on lui reconnaît le potentiel suivant :

- clarifier et cautionner l'entente entre les intervenants à propos des normes de performance visées, cette compréhension partagée étant jugée cruciale ;
- faciliter l'étalonnage entre les ministères, d'une part, et entre les secteurs public et privé, d'autre part, ce qui aide le gouvernement à choisir des fournisseurs à moindre coût et à réaliser des gains d'efficacité. La comptabilité d'exercice devrait faciliter considérablement une telle approche ;
- encourager les ministères à comprendre les chaînes de résultats qui sous-tendent leurs activités, c'est-à-dire les liens logiques (et, il faut le souhaiter, probants) entre les dépenses ministérielles, les tâches, les extrants et leur contribution aux résultats escomptés ;

· appuyer le gouvernement dans sa gestion de la complexité, notamment lorsque les travaux de plusieurs ministères convergent à la réalisation d'un objectif particulier (notion de « gouvernement décloisonné »).

Pendant que ce nouveau train de réformes était mis en œuvre, le gouvernement rognait la prééminence du ministère des Finances et de l'Administration dans le processus politique du cycle budgétaire. Les divisions chargées de la surveillance de chaque ministère sectoriel ont vu leur niveau de dotation et leur rôle diminuer sensiblement. Le gouvernement a, en revanche, eu davantage recours à des conseils externes, y compris du secteur privé.

### Quels ont été les résultats des réformes de la deuxième génération ?

En l'absence d'évaluation systématique de ces réformes, il n'est pas aisé d'analyser les résultats de la deuxième génération du système de l'évaluation de la performance et de la gestion. La sous-traitance, les appels d'offres concurrentiels, la contractualisation, la création de relations acheteur-fournisseur pour la prestation de services et la création des chartes de service, autant de pratiques qui vont dans le sens d'une meilleure performance mesurée explicitement et immédiatement et aussi probablement de gains d'efficacité conséquents.

Pour que ces approches fonctionnent de manière satisfaisante, il est nécessaire de bien comprendre les objectifs des programmes et les « niveaux cibles » de performance, de développer l'information de gestion et de disposer en permanence de mesure *ex ante* de la réalité de la performance. Bien que ces priorités soient connues, le vérificateur général a signalé les difficultés rencontrées par nombre d'organismes dans la gestion des contrats, en particulier le manque de procédures pour un examen régulier

de la performance des sous-traitants, des systèmes d'information de gestion inadéquats et des difficultés dans le choix d'indicateurs de performance appropriés (Barrett, 2001).

L'ANAO a fait des évaluations de la performance à partir d'un échantillon de rapports des ministères et des organismes remis au Parlement (déclarations budgétaires *ex ante* des cibles de dépenses et de performance du gouvernement et rapports annuels *ex post* de la performance effective des ministères) (ANAO, 2003). La vérification de performance de 2003 a permis de constater que :

- les rapports définissaient habituellement les résultats à atteindre (et leurs incidences prévues) de façon claire ;
- les organismes ont fourni de l'information exacte sur la performance, mais la plupart n'avaient pas les normes et les procédures nécessaires pour garantir la qualité des données et leur cohérence<sup>4</sup> ;
- l'information sur la performance était insuffisamment structurée pour permettre une évaluation de l'efficacité opérationnelle ou de la rentabilité des extrants effectivement produits ;
- les cibles étaient fréquemment absentes et l'étalement peu utilisé ;
- les rapports annuels donnaient seulement une information descriptive, sans analyse sur le fond, d'où la difficulté de savoir si la performance avait répondu aux attentes. La plupart de ces rapports ne s'attardaient pas sur les cas où la performance était faible ;
- les organismes ont continué d'avoir des difficultés à mesurer la qualité de leurs extrants, leur efficacité ou leur impact ;
- l'information sur la performance était minime,

<sup>4</sup> Si l'information sur la performance devient un jour importante aux fins de reddition de comptes, cette lacune sera alors rédhibitoire. La raison en est que pour jouer le jeu du système, on trouve des palliatifs à travers de l'information inexacte ou trompeuse sur la performance. Cette situation s'est produite aux États-Unis avec le programme d'éducation du gouvernement *No Child Left Behind*, qui lie le financement du gouvernement fédéral à la performance de l'école (*The Washington Post*, le 8 novembre 2003). C'est là un élément très pertinent pour la décision du gouvernement australien de lier les 5 milliards de dollars US de financement scolaire à la performance académique des étudiants (*The Sydney Morning Herald*, le 14 novembre 2003).

voire inexistante, en ce qui concerne les contributions spécifiques des ministères à la réalisation des objectifs partagés;

- les résultats des évaluations n'étaient pas toujours discutés dans le rapport annuel.

L'ANAO a conclu de cette étude que l'information sur la performance n'avait pas été présentée et analysée dans les rapports annuels d'une façon qui permettait aux parlementaires et aux autres intervenants de bien comprendre les résultats et de les interpréter. Aussi, il n'est pas surprenant que le vérificateur général ait remarqué<sup>5</sup> l'insatisfaction des comités parlementaires en ce qui concerne l'information sur la performance fournie par les ministères, à quoi s'ajoutait la difficulté représentée par les changements incessants de définition des concepts. Analyser la performance ministérielle dans le temps et faire des comparaisons d'efficacité entre les ministères étaient devenus de plus en plus difficile.

Ces écueils montrent que le gouvernement n'a pas réussi à atteindre l'un des objectifs de sa réforme, à savoir une plus grande transparence par le biais des évaluations comparatives des coûts unitaires pour tous les ministères et les organismes. L'un des avantages de l'approche systémique et centralisée en matière de mesure de la performance tient au fait qu'elle permet justement de faire de telles comparaisons. Une approche décentralisée, à l'inverse, compromet cette démarche comparative.

L'information sur la performance ne permet pas, à elle seule, de comprendre utilement la performance des programmes et des organismes. Les simples relevés de données sont en général insuffisants. Une information qualitative, par interprétation ou analyse, est nécessaire. Comme exemple de ce propos, on fera référence à des comparaisons de performance publiées tous les ans depuis 1995 entre l'action du palier fédéral et celle des États dans une vaste gamme de services gouvernementaux (écoles, hôpitaux publics, police, logements publics, soutien aux personnes âgées, etc.) (SCRCSSP, 2003). Les enquêtes montrent que, pour les ministères (fédéraux et étatiques), ces données sont utiles pour

l'analyse et la formulation de politiques, pour l'évaluation des besoins en ressources des organismes et pour les exposés faits aux ministres et aux parlements.

Compte tenu des limites de la connaissance retirée de l'information sur la performance, il est prématuré de porter un jugement global sur la nature et la qualité de l'activité d'évaluation en cours dans le gouvernement australien. En complément, des évaluations en profondeur sont nécessaires, particulièrement quand on considère les questions de causalité.

Certains ministères australiens accordent toujours une haute priorité à l'évaluation. D'autres cultivent de bonnes pratiques à cet égard (Mackay, 2004). Toutefois, il semble que le nombre des évaluations a tendance à diminuer, les ministères s'étant affranchis du caractère périodique et systématique des évaluations que préconisaient les réformes de la première génération.

Le ministère des Services à la famille et à la collectivité donne une illustration de bonne pratique en évaluation. Ses affectations budgétaires annuelles (soutien aux personnes handicapées, allocation pour le logement) se chiffrent à quelques 35 milliards de dollars américains. Le Ministère rapporte un large éventail d'information de performance (y compris sur les résultats intermédiaires et les conclusions d'évaluation) au Parlement, en particulier en ce qui concerne la quantité, la qualité et le coût des extrants. (FaCS, 2002). Le Ministère prépare régulièrement un plan de recherche et d'évaluation, dont la dernière version inclut 145 projets. Le budget total pour ce travail s'élève à 20 millions de dollars américains par an, soit l'équivalent de 1,4 % du budget opérationnel du Ministère. Six instituts universitaires sont subventionnés à cet effet pour des recherches axées sur les résultats en terme de comportements, ce qui est une dimension importante des évaluations.

Ainsi, dans un contexte de déréglementation, le ministère des Services à la famille et à la collectivité s'est donc approprié l'évaluation et lui a conféré un caractère prioritaire. Sans financement particulier

<sup>5</sup> Exposé à un séminaire de la Banque mondiale sur les *Australian Public Sector Financial Management Reforms*, le 25 juin 2003.



pour ce travail, il a cependant choisi de le financer de son propre chef, à même son enveloppe budgétaire.

L'autre question critique est celle de l'utilisation des conclusions des évaluations dans le cycle budgétaire annuel. Le ministère des Finances et de l'Administration n'a pas mené d'enquête à ce sujet et il est donc impossible de tirer des enseignements définitifs. L'effort consacré à l'évaluation par plusieurs ministères produit une grande quantité de conclusions d'évaluation qui peuvent être utilement exploitées dans les soumissions budgétaires des ministères au Cabinet.

Toutefois, Campbell (2001) a démontré que le processus de formulation de politiques s'est sensiblement édulcoré au cours des années récentes d'une part du fait du caractère irrégulier des réunions du Cabinet et de leur impréparation et, d'autre part, parce que le ministère des Finances et de l'Administration n'est plus le levier du Cabinet en ce qui concerne la performance des programmes gouvernementaux. Ces constats sont basés sur un certain nombre d'entrevues avec de hauts fonctionnaires menées à la fin des années 1990. Il en tire la leçon que le processus a été miné à cause de la tendance du gouvernement conservateur à fonctionner de façon moins collégiale que son prédécesseur travailliste. Il fait moins confiance, n'invite plus les hauts fonctionnaires à la prise de décisions et privilégie une technique budgétaire en apparence axée sur les extrants – et souvent même sur les intrants – au détriment de l'attention à prêter aux résultats.

Un haut fonctionnaire interrogé par Campbell a déploré que [traduction libre] «...si le ministère des Finances continue, comme il le fait, à ne plus participer et à ne plus avoir accès au Cabinet, alors les débats ne porteront plus sur les questions de politique et vous perdrez de l'expertise, tant celle des fonctionnaires que celle des ministres.» La véracité de cette prédiction a été confirmée récemment par une étude selon laquelle la portée de la participation du ministère des Finances et de l'Administration au processus budgétaire, la qualité de ses conseils et la fiabilité des preuves sur lesquelles s'appuient ces conseils ont connu une diminution marquée pendant la période de la réforme de

la deuxième génération. Ceci a permis à l'Organisation de coopération et de développement économique (OCDE) d'affirmer :

**« En Australie, la déréglementation des services publics et l'adoption par les organes centraux d'une position d'indépendance s'est traduite par une liberté sur le plan de la gestion; mais on considère actuellement qu'elle a privé les fonctionnaires du ministère des Finances des informations qui leurs sont nécessaires pour conseiller comme il convient leur ministre. » (OCDE, 2002)**

Pour un ministère des Finances, qui joue traditionnellement un rôle de premier plan dans le budget et dans les processus de formulation des politiques du gouvernement, cette critique est majeure. Il semble donc que l'approche décentralisée en matière d'évaluation ait contribué à réduire la capacité des ministères centraux à conseiller le gouvernement de façon adéquate.

## ■ COMPARAISON ENTRE LES RÉFORMES DE LA PREMIÈRE ET DE LA DEUXIÈME GÉNÉRATION

Il est délicat d'émettre un jugement équilibré sur le succès relatif des deux générations de réformes de l'évaluation de la performance et de la gestion en Australie. La première génération a commencé en 1987. Elle a mis l'accent sur la planification, la conduite et l'utilisation des évaluations, en particulier pendant le processus budgétaire. On a constaté qu'elle a acquis une influence importante sur les conseils de politiques et les processus budgétaires dès le milieu des années 1990. Il a été également démontré que les évaluations étaient utilisées de façon intensive par les ministères dans leur gestion courante. Toutefois, la faiblesse de cette première génération tenait au manque d'attention accordé à la collecte, au rapport et à l'utilisation de l'information sur la performance dans les programmes.

Les réformes de la deuxième génération sont en cours depuis huit ans. Elles s'ancrent sur une approche beaucoup plus décentralisée en matière de gestion du secteur public. Avec elles, on vise à améliorer la performance du secteur public en actionnant le levier d'incitatifs offerts aux PDG et à leurs ministères. Toutefois, l'effet de ces incitatifs sur la qualité de la performance n'est pas manifeste. Dans une certaine

mesure, l'information sur la performance demeure insuffisante pour une bonne gestion gouvernementale et ne répond pas aux besoins du Parlement.

Certains ministères persistent à pratiquer l'évaluation à grande échelle à l'appui de leur gestion interne et aux fins de rapports parlementaires, ce qui peut être interprété, en partie, comme un avantage résiduel des réformes de la première génération. Pour autant, on ignore dans quelle mesure les conclusions d'évaluation servent encore aux ministères dans le processus budgétaire annuel. Enfin, la capacité du ministère des Finances et de l'Administration de fournir de bons conseils de politique semble avoir été considérablement restreinte.

L'élément central du système de gestion de la performance du gouvernement actuel tient au « cadre des résultats et des extrants ». La timidité des incitatifs semble être la faille de ce cadre. Les redditions de comptes qui le soutiennent sont instituées pour l'essentiel entre les services et leurs ministres d'une part, et le Parlement d'autre part. La reddition de comptes des ministères sectoriels auprès du ministère des Finances et de l'Administration et des autres ministères centraux est passée au second plan.

La reddition de comptes au Parlement a toujours été un élément fondamental de la gouvernance. Or, de toute évidence, le Parlement continue à examiner de près les budgets, et d'abord par tradition les intrants, c'est-à-dire les affectations de dépenses. Le Parlement n'a pas endossé le nouveau rôle qu'il était censé jouer. Certes l'ANAO rapporte directement au Parlement, mais son témoignage ne saurait remplacer la richesse de l'échange entre les ministères et les parlementaires qu'avait instauré la première génération de réformes.

La première génération de réformes visait à trouver un point d'équilibre entre une approche décentralisée du style « laisser les gestionnaires gérer » et les exigences formelles du gouvernement de rapports sur la performance en appui à la prise de décisions budgétaires et à la reddition de comptes. La deuxième génération de réformes a appliqué une approche décentralisée au système d'évaluation avec, pour conséquence dommageable, que la qualité et la profondeur de l'information accessible sur la per-

formance gouvernementale ont effectivement diminué. Certaines erreurs des années 1980 ont été répétées, à la différence que les ministères ont désormais l'obligation de produire un volume important d'information sur la performance et de la publier. Ainsi, les réformes de la seconde génération n'ont-elles pas détruit l'édifice mis en place.

## ■ ENSEIGNEMENTS EN VUE DE L'ÉDIFICATION D'UN SYSTÈME DE PERFORMANCE

Dans la présente section, nous souhaitons tirer les leçons de l'expérience australienne présentée dans les sections antérieures et des politiques des autres pays qui se sont efforcés de consolider ou d'établir un système d'évaluation pour l'ensemble du gouvernement (Mayne et Zapico-Goni, 1997; Mackay, 2005; Boyle et Lemaire, 1999).

En vertu de la première génération de réformes, la stratégie d'évaluation australienne était une initiative centralisée du tout puissant ministère des Finances, étroitement liée à son principal domaine d'influence, le processus budgétaire annuel. Le succès de la stratégie peut être raisonnablement imputé au pouvoir de ce ministère. Il existe des avantages et des inconvénients à toute approche centralisée. Elle provoque souvent un faible engagement de la part des ministères sectoriels. Le ministère des Finances s'est efforcé de pallier cet inconvénient à la fois par la persuasion et en donnant du soutien

Les limites du recours aux bonnes pratiques apparaissent avec la deuxième génération de réformes (et les effets initiaux des réformes du style « laisser les gestionnaires gérer » au milieu des années 1980). La combinaison de « carottes », de « bâtons » et de « sermons » semble en fin de compte plus efficace que de faire appel à un seul de ces moyens.

Dans le cas de l'Australie, les « carottes » avaient trait au soutien du ministère des Finances et donc à la promesse d'ententes pour les ministères sectoriels qui démontraient de façon convaincante la nécessité d'un financement direct pour des programmes non performants. Les « bâtons » ressortaient du lien étroit du ministère des Finances avec les ministères et de sa capacité à influencer sur leurs affectations bud-

gétaires. Le ministère des Finances avait aussi la possibilité d'embarrasser les ministères en publiant les classements comparatifs qu'il préparait sur les approches des ministères en matière de planification et de conduite d'évaluation. Les « sermons » se manifestaient enfin à travers les invocations du secrétaire permanent du ministère des Finances et de ses hauts fonctionnaires.

Les ministères des finances n'accueillent pas toujours favorablement des réformes qui accroissent la quantité d'information disponible sur la performance gouvernementale. Avant les réformes, le ministère des Finances de l'Australie participait activement à l'examen détaillé des dépenses des ministères. Cette insistance traditionnelle sur les dépenses recèle un danger car on prêtera moins d'attention aux résultats. Les ministères des finances peuvent même faire obstacle à la réforme. Le fait que le ministère des Finances australien soit chargé de la supervision du système d'évaluation assurait sa « primauté » sur les autres ministères.

Le changement de culture au sein du ministère des Finances a d'ailleurs résulté d'un processus lent, étalé sur un certain nombre d'années, qui s'est accompagné d'une rotation importante du personnel. L'insistance nouvelle du Ministère sur la rentabilité (plutôt que sur la dépense) a eu des répercussions sur la nature et la qualité des conseils que le Ministère donnait au Cabinet, ces conseils s'appuyant de plus en plus sur les conclusions d'évaluation.

L'expérience australienne prouve que la discussion budgétaire annuelle offre la meilleure opportunité de revenir sur la performance du gouvernement. Car se posent alors des questions essentielles : quels sont les critères de qualité d'un conseil sur une politique ? Quelles sont les normes appliquées dans les ministères et par le Cabinet<sup>6</sup> ? Quels mécanismes consultatifs rejoignent le mieux les besoins d'information du cabinet ? Quelles sont les compétences et les qualifications exigibles des fonctionnaires qui œuvrent comme analystes de politiques ?

Une autre leçon tient à l'importance de confier le leadership des réformes à un organisme désireux de s'en faire le champion et d'investir du temps et de l'énergie pour qu'elles réussissent. En Australie avec la deuxième génération de réformes, mais aussi au Canada, la simple promotion du caractère utilitaire de l'évaluation a échoué. Les interventions de l'ANAO ou du Bureau du Premier ministre ont été en revanche déterminantes.

Comme tous les autres volets d'une réforme gouvernementale, le renforcement des systèmes d'évaluation et d'information sur la performance exige un long délai de mise en œuvre et un suivi permanent. Les bienfaits et les limites d'une réforme apparaissent seulement de façon évidente au bout de plusieurs années et après maintes réorientations.

L'étroitesse de la base sur laquelle elle prend appui représente le péril le plus grave pour une réforme. Si elle dépend du bon vouloir d'une personne ou d'un groupe restreint de personnes et que ces personnes changent d'emploi, le mouvement réformiste risque de péricliter. Il faut donc le plus rapidement possible institutionnaliser la réforme, d'autant qu'on connaît l'effort considérable à consentir avant que l'engagement pour la réforme pénètre les consciences.

Avec les deux générations australiennes de réformes, les exigences formelles et les principes de l'information sur la performance et des évaluations étaient moins prépondérants que leur mode de mise en œuvre. Il est sans aucun doute utile, mais insuffisant, d'avoir une décision du Cabinet (ou dans les pays à tradition napoléonienne, une loi ou un décret) pour habiliter le changement. Il est beaucoup plus important de proposer des incitatifs de nature à produire cette information puis à l'utiliser dans la perspective des objectifs désignés (prise de décisions budgétaires, bonne gestion ou reddition de comptes). En vérifiant qu'il y a une forte demande pour cette information, on est sûr de réussir. D'où la nécessité de bien détecter qui a besoin de quelle information, quand et à quelles fins.

<sup>6</sup> Les caractéristiques des bons conseils de politique et le rôle que peut jouer l'évaluation pour les appuyer au gouvernement fédéral sont étudiés par Uhr et Mackay (1996).

Il est faux de résumer la mise en œuvre d'une réforme du secteur public à la résolution d'une banale question technique. Les besoins d'information des organismes centraux ou du Parlement peuvent être satisfaits si les ministères sont fortement incités à fournir cette information. Mais l'intégration d'une réforme dans les processus de planification et de gestion d'un ministère, les cas échéant en les bouleversant, implique un changement radical dans la culture du ministère. Un tel virage a été négocié dans le ministère australien des Finances en raison du fort niveau d'engagement de deux secrétaires ministériels successifs. En tout état de cause, il n'existe pas de recette miracle pour aboutir à un changement de culture, que ce soit au sein d'un ministère ou dans l'ensemble de la fonction publique.

L'autre erreur fréquente dans la conception d'un système d'évaluation de la performance consiste à vouloir téléguider l'information, tant qualitative que quantitative. Cette erreur survient si on conçoit un système d'indicateurs de performance trop global ou si on commande un nombre élevé d'évaluations complexes et coûteuses. Un tel gaspillage a été observé dans plusieurs pays.

Un dernier débat impossible à trancher est celui de l'objectivité et de l'indépendance de l'évaluation. Les évaluations externes sont considérées comme objectives, mais elles pâtiennent d'une mauvaise connaissance des programmes. En outre, le niveau d'engagement de l'organisme évalué est souvent faible, ce qui peut entraîner une utilisation interne erronée des conclusions d'évaluation. À l'inverse, les autoévaluations sont décriées comme étant égo-centriques et manquant d'objectivité et de rigueur. Dans les ministères australiens, les évaluations sont souvent effectuées par des unités spécialisées du ministère, indépendantes de l'aire hiérarchique responsable de la gestion du programme évalué. Les évaluations importantes sont supervisées par les comités directeurs qui font appel à des représentants d'autres ministères sur lesquels on s'appuie pour attester de l'objectivité et de la rigueur de la démarche et pour veiller à ce que les irritants ne soient pas mis à l'écart de l'évaluation. Cette approche est aussi le gage d'un fort niveau d'engagement et de l'utilisation des conclusions par les intervenants de l'évaluation.

## BIBLIOGRAPHIE

- ANAO (Australian National Audit Office) (1991). *Evaluation in Preparation of the Budget*, Efficiency Audit Report N° 13, 1991-92, Canberra: AGPS.
- ANAO (Australian National Audit Office) (1993). *Program Evaluation – Strategies, Practices and Impacts – Industry, Technology and Regional Development Portfolio*, Efficiency Audit Report N° 35, 1992-93, Canberra: AGPS.
- ANAO (Australian National Audit Office) (1997). *Program Evaluation in the Australian Public Service*, Performance Audit Report N° 3, 1997-98, Canberra: AGPS.
- ANAO (Australian National Audit Office) (2003). *Annual Performance Reporting*, Performance Audit Report N° 11, 2003-04, Canberra: AGPS.
- Barrett, P., Auditor-General of Australia (2001). *Evaluation and performance auditing: sharing the common ground. A review of developments*, Address to the Australasian Evaluation Society, Canberra, le 10 octobre.
- Boyle, R. et D. Lemaire (1999). *Building Effective Evaluation Capacity: Lessons from Practice*, New Brunswick and London, Transaction Publishers.
- Campbell, C. (2001). Juggling Inputs, Outputs, and Outcomes in the Search for Policy Competence: recent experience in Australia. *Governance: An International Journal of Policy and Administration*, 14 (2), pp. 253-282.
- Crossfield, L. et A. Byrne (1994). *Review of the Evaluation Function in DEET*, Canberra: Department of Employment, Education and Training, and Department of Finance.
- DoF (Department of Finance) (1994). *The Use of Evaluation in the 1994-95 Budget*, Finance Discussion Paper, Department of Finance, Canberra.

- DoFA (Department of Finance and Administration) (2000). *The Outcomes & Outputs Framework: Guidance Document*, Canberra: DoFA (Department of Finance and Administration) (2003). *Annual Report 2002-2003*, Canberra: AGPS.
- Edwards, M. (2002). «Public sector governance – future issues for Australia.» *Australian Journal of Public Administration* 61 (2), p. 51-61.
- FaCS (Department of Family and Community Services) (2001). *Research and Evaluation Framework*, Canberra: FaCS.
- FaCS (Department of Family and Community Services) (2002). *Annual Report 2001-02*, Canberra: AGPS.
- Mackay, K. (1998). *The Development of Australia's Evaluation System*. ECD Working Paper N° 4. Washington D.C. : Département de l'évaluation des opérations, Banque mondiale.
- Mackay, K. (2004). *Two Generations of Performance Evaluation and Management System in Australia*. ECD Working Paper no11, Washington D.C., Operations Evaluation Department, The World Bank.
- Mackay, K. (2005). *Institutionalization of monitoring and evaluation systems to improve public sector management*, Article présenté à la conférence inaugurale de l'International Development Evaluation Association, New Delhi.
- Mayne, J. et E. Zapico-Goni (eds.) (1997). *Monitoring Performance in the Public Sector: Future Directions from International Experience*, New Brunswick and London: Transaction Publishers.
- OCDE (Organisation de coopération et de développement économiques) (2002). *Modernisation du secteur public: un nouveau programme*, GOV/PUMA (2002) Paris: OCDE.
- Russell, D. (2003). «Performance management and evaluation: the Australian experience», dans Jay-Hyung Kim (ed.), *Developing a Performance Evaluation System in Korea*, Seoul: Korea Development Institute.
- Schick, A. (2001). *Does Budgeting Have a Future?*, OECD (PUMA/SBO (2001) Paris: OECD.
- Schick, A. (2002). «La budgétisation a-t-elle un avenir?», *Revue de l'OCDE sur la gestion budgétaire*, Paris: OCDE, vol. 2, n° 2, p. 7 à 52.
- Schick, A. (2003). *The Performing State: Reflection on an Idea Whose Time Has Come But Whose Implementation Has Not*, OCDE (GOV/PUMA/SBO (2003) Paris: OCDE.
- Schick, A. (2004). «L'État performant: réflexions sur une idée entrée dans les esprits mais pas encore dans les faits», *Revue de l'OCDE sur la gestion budgétaire*, Paris: OCDE, vol. 3, n° 2, p. 79 à 119.
- SCRCSSP (Steering Committee for the Review of Commonwealth/State Service Provision) (2003). *Report on Government Services 2003*, Canberra: AusInfo.
- The Sydney Morning Herald*. Le 14 novembre 2003. *Threat to Schools: Perform or Be Shut Down*.
- Secrétariat du Conseil du Trésor du Canada (2000). *Des résultats pour les Canadiens et les Canadiennes: Un cadre de gestion pour le gouvernement du Canada*, Ottawa: Secrétariat du Conseil du Trésor.
- Uhr, J. et K. Mackay (eds.) (1996). *Evaluating Policy Advice: Learning from Commonwealth Experience*, Canberra: Federalism Research Centre of the Australian National University and DoF.
- The Washington Post*. Le 8 novembre 2003. *Education «Miracle» Has a Math Problem*.

# Repères et Références

Télescope a choisi des articles et des ouvrages à lire utilement en complément de ce numéro consacré à l'évaluation de politiques et de programmes publics.

## ■ DES ARTICLES

### An Evaluation Approach to Public Policy

Par **Andrea Simpson** (2005)

Langue : anglais

À consulter : <http://www.optimumonline.ca/print.phtml?id=234>

Comment associer la prise de décision à haut risque, l'obligation de transparence et le souci de toujours travailler dans le but d'offrir le meilleur service à la population? L'auteur fait de cette mise en tension inévitable le défi ultime de l'évaluation. L'article ébauche un modèle préliminaire de ce que pourrait être une approche évaluative d'élaboration et d'implantation des politiques publiques.

### Revue canadienne d'évaluation de programme

Par la **Société canadienne d'évaluation**

vol. 20, n° 2, automne 2005

Langue : français et anglais

À consulter en version imprimée

à la bibliothèque de l'ENAP à Québec

Dans le numéro d'automne 2005 de la *Revue canadienne d'évaluation de programme*, une section est consacrée au thème des compétences de l'évaluateur. Six articles sont publiés sur le sujet. Ils traitent de paramètres tels que l'augmentation de la performance, les programmes de formation, la reconnaissance ou parfois l'absence de reconnaissance de la fonction d'évaluateur... Parmi les plus intéressants, l'un des articles révèle les pratiques expérimentées au Japon pour élaborer et instaurer un système d'agrément et de reconnaissance professionnelle dans le domaine de l'évaluation.

### Améliorer le professionnalisme de l'évaluation

Par **T.K. Gussman Associates Inc.** (2005)

Rapport préparé pour le Centre d'excellence en évaluation, Secrétariat du Conseil du Trésor, Gouvernement du Canada

Langue : français

À consulter : [http://www.tbs-sct.gc.ca/eval/dev/Professionalism/profession\\_f.asp](http://www.tbs-sct.gc.ca/eval/dev/Professionalism/profession_f.asp)

Le présent rapport s'attache à décrire les méthodes visant à accroître le professionnalisme de la fonction d'évaluation au gouvernement fédéral. Selon ses auteurs, le défi de la crédibilité, pour les professionnels de l'évaluation, repose sur l'amélioration de l'image de leur fonction ainsi que sur l'assurance de garantir une utilisation optimale des évaluations, par les gestionnaires et les administrateurs, dans l'optique d'une prise de décision adéquate.

### Political Institutions, Policymaking Processes and Policy Outcomes in Chile

Par **Cristóbal Aninat, John Landregan, Patricio Navia et Joaquín Vial**

de l'Inter-American Development Bank (2006)

Langue : anglais

À consulter : <http://www.iadb.org/res/publications/pubfiles/pubR-521.pdf>

Dans cet article, les auteurs tentent de mesurer l'influence des institutions politiques sur le processus d'élaboration des politiques publiques au Chili. Ils traitent par ailleurs des liens de cause à effet qui peuvent exister entre l'élaboration ou la poursuite des politiques publiques et les résultats constatés et découlant de l'implantation de celles-ci. Pour ce faire, les auteurs s'appuient sur le cadre théorique développé par Spiller et Tommasi.

### Politiques publiques et santé : les dilemmes de l'évaluation prospective

Par **Jean Turgeon, France Gagnon, Jacques Bourgault et Sophie Garant**

Administration publique du Canada, vol. 48, n° 3, automne 2005, p. 328-347

Langue : français

À consulter en version imprimée à la bibliothèque de l'ENAP à Québec

Les conséquences des actions gouvernementales sur la santé de la population sont de plus en plus prises en considération par les gouvernements. En vertu de l'article 54 de la *Loi sur la santé publique*, les ministères sont tenus, au Québec, de mesurer les différents impacts (sociaux, environnementaux et économiques) de leurs actions sur la santé. Pour atteindre cet objectif, les auteurs de l'article suggèrent d'introduire un processus d'évaluation prospective des politiques publiques. Ils ne taisent pas les difficultés que soulève un tel changement et concluent en soulignant que la mise en œuvre et la réussite de ce type d'évaluation doivent être soutenues et accompagnées par un leadership politico-administratif fort.

### Private Governments or Public Policy Tools? The Law and Public Policy of New Jersey's Special Improvement Districts

Par **Jonathan B. Justice et Robert S. Goldsmith**

Dans l'*International Journal of Public Administration*, vol. 29, n° 1-3, 2006, p. 107-139

Langue : anglais

À consulter en version électronique par la banque de données Metapress

Cet article fait le point sur la création et l'évolution des *Business Improvement Districts* (BID). Implantés au New Jersey,

les BID sont des zones spécifiques gérées par des entreprises et par des propriétaires fonciers. L'auteur a analysé plusieurs exemples de BID. Il soutient qu'il s'agit d'authentiques partenariats public-privé qui sont mis en œuvre à la fois comme instruments de politiques publiques, défendant l'intérêt public, et comme entités autonomes (*self-help entities*), lesquelles visent des intérêts particuliers.

## ■ DES OUVRAGES

### Politiques publiques comparées dans les États fédérés : l'Allemagne, l'Australie, le Canada, les États-Unis et la Suisse

Sous la direction de **Louis Imbeau** (2005)  
Langue : français  
Éditeur : Les Presses de l'Université Laval

Dans cet ouvrage, les auteurs avancent une théorie selon laquelle les États fédérés détiennent un degré d'autonomie suffisant pour leur permettre d'élaborer, d'approuver et de déployer des politiques publiques originales. Ceci a pour conséquence de faciliter l'analyse comparative des politiques publiques. À partir des exemples des États fédérés de cinq pays (länder allemands, États australiens, provinces canadiennes, États américains et cantons helvétiques), les chercheurs analysent et mettent en perspective les actions menées dans divers secteurs tels que la santé, l'éducation, l'environnement ou les finances publiques. Cet exercice doit contribuer, selon eux, à la consolidation des fondements théoriques et à l'avancée des connaissances dans le domaine de l'analyse des politiques publiques.

### American Public Policy: Promise and Performance

Par **B. Guy Peters** (2004)  
Langue : anglais  
Éditeur : Seven Bridges Press, LLC

La sixième édition de cet ouvrage permet à l'auteur de mettre à jour l'information relative au contenu et aux procédures d'implantation des politiques publiques. Ainsi, les politiques dans les domaines – entre autres – des soins de santé, de l'éducation, de la sécurité sociale, de l'énergie, de l'environnement ou encore de la défense nationale font l'objet d'un réexamen. Cette nouvelle édition contient également des ajouts dans lesquels sont étudiés les impacts des événements du 11 septembre 2001, notamment la réapparition des déficits budgétaires dans l'élaboration des politiques publiques aux États-Unis. En outre, l'auteur analyse les actions menées pour implanter des législations basées sur les résultats (*results-based legislation*) comme la *No Child Left Behind Act*. Il s'attarde enfin, dans un tout nouveau chapitre, à ausculter les politiques sociales et les conflits interethniques.

### Institutionnaliser l'évaluation des politiques publiques : étude comparée des dispositifs en Belgique, en France, en Suisse et aux Pays-Bas

Par **Steve Jacob** (2005)  
Langue : français  
Éditeur : Peter Lang

L'évaluation des politiques des institutions étatiques jette un regard objectif sur la conduite de l'action publique. Son but est d'informer les décideurs politico-administratifs et les citoyens intéressés au fonctionnement de l'État, sur les effets induits par la mise en œuvre de ces politiques. À travers une comparaison détaillée de situations observées dans quatre pays, cet ouvrage présente les configurations institutionnelles mises en place pour favoriser l'essor de la pratique évaluative. La compréhension des facteurs explicatifs des différences entre les cas étudiés se forge, d'une part, sur les principales singularités des systèmes politico-administratifs et, d'autre part, sur des éléments de nature conjoncturelle comme la lutte aux déficits publics ou la volonté de renforcer la légitimité des institutions publiques. En plus d'offrir une des premières analyses comparées en la matière, cet ouvrage éclaire, sous un angle nouveau, les théories « institutionnalistes » et d'analyse des organisations en s'appuyant sur une recherche empirique approfondie.

# Changer les organisations publiques de l'intérieur.

## Le cas de l'évaluation à la Banque mondiale

Par **Mita Marra**, Conseil national italien de la recherche – Institut pour l'étude des sociétés méditerranéennes, Naples, Italie.  
Consultante auprès de la Banque mondiale • Mita.Marra@issm.cnr.it  
Traduit de l'anglais

Au niveau mondial, les récentes stratégies de réforme du secteur public ont largement insisté sur l'évaluation de politiques et de programmes comme un élément central du processus de réforme (OCDE, 1995). L'évaluation systématique des politiques et des programmes afin de mesurer leurs résultats et d'améliorer leurs performances futures (Weiss, 1997) est actuellement vue comme faisant partie intégrante de « la gestion axée sur les résultats » – une des phrases clés de la réforme du secteur public (Boyle, Lemaire, 1999). Selon Chelimsky et Shadish (1997), le contexte politique et administratif dans lequel les évaluations sont effectuées est en train de changer. Le mouvement mondial – tant dans les pays industrialisés que ceux en développement – va vers une réduction des impôts et des taxes, des déficits, des transferts financiers des riches vers les pauvres et de la taille de l'État. Chelimsky et Shadish (1997) remarquent que « du fait de ce mouvement, l'évaluation prend une grande importance avec le besoin de plus en plus ressenti de mesurer l'efficacité et le coût des politiques et des programmes. » L'évaluation devient ainsi un outil unique et précieux dans de nombreux domaines du secteur public et particulièrement dans le domaine de l'aide au développement qui doit tenir compte des nouveaux besoins économiques, sociaux et politiques (Chelimsky et Shadish, 1997).

Dans le domaine de l'aide au développement, comme dans celui de la vérification et de la mesure de la performance des administrations publiques, le décideur voit l'évaluation comme un moyen d'améliorer la politique menée et de répondre aux exigences de reddition de comptes. Le récent rapport de la commission Meltzer du Sénat américain, qui examine les effets de l'aide au développement offerte par la Banque mondiale et d'autres agences de développement régional, insiste fortement sur la nécessité de renforcer la prise de conscience envers les résultats et la reddition de comptes afin d'améliorer l'efficacité de l'aide (International Financial Institutions Advisory Commission, 1998). Svensson (1997) note que la publication des résultats occupe actuellement une place prépondérante dans la politique d'aide au développement: « Une diffusion ouverte et honnête des résultats est nécessaire pour qu'un donateur soit motivé pour poursuivre ses efforts sur le long terme ».

Des efforts considérables ont été consentis dans les organismes publics et internationaux pour institutionnaliser l'évaluation au point d'en faire une fonction systématique. Or, bien que des ressources humaines et techniques soient disponibles à l'intérieur même du système public, rien ne garantit que l'évaluation, qu'elle soit interne ou externe, ait un impact significatif sur la pratique et la mise en œuvre des politiques (Thoenig, 1999).

En premier lieu, l'évaluation, comme le budget et l'audit, n'est pas qu'un concept théorique. Sa pratique reflète les différences de cultures professionnelles et politiques, car des communautés diverses développent cette fonction dans des contextes politiques éloignés (Gray, Jenkins et al., 1993). En second lieu, la principale préoccupation concernant l'évaluation est liée aux changements dans les politiques et les



pratiques institutionnelles. Les arrangements institutionnels formels, les procédures et les pratiques coutumières forment un cadre à travers lequel l'évaluation influence la prise de décision.

Dans une perspective organisationnelle, cet article traite du rôle de l'évaluation vis-à-vis des réformes de gestion récemment introduites, ainsi que des raisons de son utilisation. Basé sur quatre études de cas et des entrevues avec des gestionnaires de la Banque mondiale, cet article examine comment l'évaluation, de la manière dont elle a été institutionnalisée à la Banque mondiale, est intégrée à la prise de décision dans la gestion des programmes de développement.

Notre objectif est de rendre compte ici des possibilités et des contraintes institutionnelles et organisationnelles qui peuvent favoriser ou gêner l'utilisation de l'évaluation pour améliorer la performance organisationnelle. Nous nous concentrerons sur trois thèmes qui sont apparus comme essentiels pour caractériser l'influence de l'information évaluative sur la prise de décision à la Banque mondiale, soit :

- le modèle institutionnel spécifique de l'évaluation et son rôle de « contrôle et équilibre » vis-à-vis de la professionnalisation de toutes les activités des employés pour une gestion axée sur les résultats ;
- la culture bureaucratique, les valeurs et les idéologies de l'organisation qui encadrent l'étude évaluative et son acceptation par les gestionnaires ;
- le modèle de consensus politique autour des recommandations faites par l'évaluation afin de promouvoir un changement de politique et de méthode de travail.

Ces trois sujets sont pertinents pour comprendre comment l'évaluation peut contribuer à l'amélioration de la performance des organisations, lorsque les politiques, les procédures institutionnelles, les règles informelles, les normes et les valeurs contraignent le comportement des gestionnaires de programmes et des décideurs. La première section de l'article décrit la méthodologie de l'étude, en clarifiant les critères de sélection des études de cas et la procédure de sélection des interviewés. La deuxième section détaille le schéma institutionnel d'évaluation ainsi que les processus de travail et de production à la Banque mondiale. La troisième section traite de

la contribution de l'évaluation à la gouvernance démocratique et institutionnelle en la rapprochant de la gestion axée sur les résultats.

La quatrième section souligne comment la culture bureaucratique et les valeurs de l'organisation imprègnent les attentes et les comportements organisationnels, affaiblissant la crédibilité de l'étude évaluative et son potentiel de changement. La cinquième section explore les différentes manières par lesquelles les gestionnaires et les décideurs acceptent progressivement les recommandations de l'évaluation et les prennent en compte, consciemment ou inconsciemment, dans des actions et des décisions subséquentes. Enfin, la sixième section tire un certain nombre de conclusions et d'implications pour une plus grande performance organisationnelle grâce à une relation améliorée entre l'évaluation, le management et les structures de gouvernance.

## ■ MÉTHODOLOGIE

Afin de démontrer comment l'information basée sur l'évaluation est intégrée dans la prise de décision organisationnelle, nous avons choisi la méthode de l'étude de cas en nous concentrant sur les quatre rapports d'évaluation suivants : (1) l'examen de la stratégie forestière de la Banque mondiale de 1991 ; (2) l'évaluation des grands barrages et des déplacements de populations ; (3) l'impact des Revues de dépenses publiques (PER) ; (4) l'évaluation à mi-parcours des activités de lutte anticorruption de l'Institut de la Banque mondiale (WBI). Ces cas concernent l'évaluation de programmes de la Banque mondiale dans trois domaines majeurs de l'aide au développement : le renforcement institutionnel et la mise en place de moyens pour la planification fiscale, les grandes infrastructures avec des implications économiques et sociales et la gestion et la préservation de l'environnement.

De plus, les quatre évaluations s'intéressent à différents aspects des programmes de développement : (1) l'étude forestière se concentre sur la mise en œuvre de la stratégie forestière de la Banque mondiale de 1991 ; (2) l'évaluation des grands barrages est une analyse coût-bénéfice ex post ; (3) l'évaluation des PER présente les incidences du contenu, du processus et des techniques, en exposant les Revues de dépenses publiques comme un outil

d'analyse; (4) l'évaluation à mi-parcours sur la lutte anticorruption est une étude théorique considérant à la fois les mises en œuvre et les résultats des premières initiatives anticorruption.

La variété des cas a été recherchée afin de détecter des schémas communs et de fournir un aperçu des types d'information que les décideurs perçoivent comme étant plus ou moins utiles (voir tableau 1). Entre janvier 1999 et mai 2001, les données nécessaires à cette recherche ont été recueillies grâce à: (1) des entretiens semi-structurés auprès de quatre catégories de répondants à la Banque mondiale; (2) des observations participatives; (3) des analyses

de contenus de documents officiels et de rapports d'évaluation.

Nous avons interrogé quatre catégories différentes de répondants. Ces catégories incluaient des évaluateurs, des décideurs de haut niveau, des gestionnaires de programmes-pays et des experts. En premier, les évaluateurs, y compris les auteurs des études évaluatives sélectionnées, ont été interrogés afin de saisir de quelle manière ils voulaient utiliser les études qu'ils ont conduites, mais aussi pour relater comment ces évaluations ont été élaborées, mises en œuvre et diffusées à l'intérieur et à l'extérieur de la Banque mondiale.

**TABEAU 1 : LES QUATRE ÉVALUATIONS**

	ÉTUDE DE FORESTERIE	REVUE DES DÉPENSES PUBLIQUES
<b>POLITIQUE/ PROGRAMME</b>	Mise en oeuvre de la stratégie forestière de 1991 de la Banque mondiale.	Procédure, contenu et technique de l'étude des dépenses publiques.
<b>RÉSULTATS INTÉRESSANTS DE LA POLITIQUE/ DU PROGRAMME</b>	Le rôle de la Banque mondiale dans le secteur forestier, tant dans les pays riches que dans les pays pauvres en forêt, et dans l'atténuation de la pauvreté à l'intérieur des objectifs de développement. La Banque mondiale soutient les expériences d'abattage sans compromettre les objectifs de conservation de l'acte d'interdiction de financer les acquisitions d'équipement aux fins d'abattage dans les forêts tropicales humides.	Qualité, opportunité, coût-efficacité et impacts sur l'aide externe et la réforme des dépenses publiques dans les pays en voie de développement.
<b>DATE DU RAPPORT FINAL</b>	2000-06-23 - FY2000	1998-11-13 - FY1999
<b>EMPLACEMENT DU SITE ÉVALUÉ</b>	Siège de la Banque mondiale, Brésil, Chine, Cameroun, Costa Rica, Inde et Indonésie	Processus d'évaluation interne.
<b>MÉTHODOLOGIE</b>	<ol style="list-style-type: none"> <li>1) Étude de cas sur le terrain;</li> <li>2) Large consultation auprès des protagonistes;</li> <li>3) Examen de la politique globale et de six portefeuilles régionaux de projets.</li> </ol>	<ol style="list-style-type: none"> <li>1) de la méthodologie d'évaluation de l'OED à un échantillon stratifié de PER par un groupe d'experts;</li> <li>2) Enquête indépendante sur les gestionnaires de projets;</li> <li>3) Entretien avec de hauts fonctionnaires de pays en voie de développement;</li> <li>4) Concertation avec des agents du Fonds Monétaire International (FMI) et d'autres partenaires.</li> </ol>
<b>PRINCIPALES CONSTATATIONS</b>	<ol style="list-style-type: none"> <li>1) La Banque mondiale a mis en œuvre la stratégie forestière de 1991 en partie seulement et pour l'essentiel par la prise en compte, dans ses prêts à caractère environnemental, d'un plus grand nombre d'éléments liés à la forêt.</li> <li>2) Les objectifs ambitieux de la stratégie sur la conservation n'ont pas été suivis d'actions suffisantes pour être atteints.</li> <li>3) L'efficacité de la stratégie a été limitée, de même que la viabilité de ses résultats.</li> </ol>	<ol style="list-style-type: none"> <li>1) Ce rapport montre que la qualité du PER a augmenté ces dernières années;</li> <li>2) Le PER a fourni des analyses des politiques de dépenses de bonne qualité mais souvent datées. Peu d'attention a été portée à l'analyse coût-efficacité ou à la qualité des services publics;</li> <li>3) Les PER n'ont qu'un impact réduit sur les stratégies de prêt de la Banque mondiale, la politique de dépenses du client et la coordination de l'aide, car les fonctionnaires du pays concerné ne perçoivent pas son utilité.</li> </ol>

Les décideurs ont été interrogés ensuite afin de savoir comment ils réagissaient à l'information découlant de l'évaluation, leur façon d'utiliser l'évaluation pour la prise de décision et leur attitude envers l'évaluation. Les gestionnaires de programmes-pays ont été rencontrés afin de comprendre quels éléments de l'évaluation étaient susceptibles de les conduire à utiliser cet outil dans l'élaboration d'un programme et dans sa mise en œuvre. Les entretiens portaient à la fois sur leurs expériences passées et sur leurs impressions *a priori* sur ces points. Enfin, les experts du secteur de l'évaluation sélectionnée (foresterie, construction de barrages, lutte contre la corruption et Revue des dépenses publiques) ont été interrogés afin de recueillir leurs commentaires sur les aspects techniques des analyses.

Le tableau 2 spécifie la composition des quatre catégories de répondants et indique le nombre d'entretiens conduits pour chacune d'entre elles.

**TABLEAU 2 : RÉPARTITION DES ENTRETIENS**

PROFIL DES RÉPONDANTS	NOMBRE D'ENTRETIENS
Évaluateurs	8
Décideurs	7
Gestionnaires de programmes	27
Experts	14
<b>Total</b>	<b>56</b>

### ÉVALUATION DES GRANDS BARRAGES

Impacts de 50 barrages, dont la construction a été financée par la Banque mondiale, sur le développement économique, social et environnemental en regard des critères et des politiques de préservation de la Banque.

Amélioration et développement de la production d'énergie, de l'irrigation et de la distribution d'eau à des fins industrielles et domestiques.

1996-08-15 - FY1997

Analyse depuis le siège de la Banque.

- 1) Analyse de contenu des documents concernant les projets de grands barrages financés par la Banque;
- 2) Questionnaire envoyé aux agences chargées de la mise en œuvre;
- 3) Entretiens informels avec le personnel opérationnel. Ces renseignements ont été étudiés afin de déterminer les projets qui ont, dans l'ensemble, réussi à aider l'atteinte de l'objectif du développement durable ainsi que les cas pour lesquels les impacts néfastes sont inacceptables.

- 1) 74 % des 50 barrages concernés par l'étude ont contribué de manière essentielle au développement économique en matière de production d'énergie, de contrôle des inondations, de fourniture d'eau aux villes et aux industries et d'irrigation;
- 2) Les réimplantations de populations ont été insuffisamment prises en charge dans la moitié des projets. En tout, les 50 barrages ont déplacé à peu près 830 000 personnes;
- 3) Les 50 barrages enregistrent des résultats mitigés en ce qui concerne la gestion des conséquences environnementales de leur activité.

### ÉVALUATION À MI-PAROURS DE LA LUTTE ANTICORRUPTION

Formation offerte en journalisme d'investigation.

Résultats dans la lutte anticorruption et le développement institutionnel.

FY1998

Tanzanie, Ouganda.

- 1) Approche fondée sur la théorie;
- 2) Études de cas sur le terrain;
- 3) Entretiens approfondis;
- 4) Analyse de contenu documentaire;
- 5) Critique empirique de la littérature en sciences sociales.

- 1) La prise de conscience menant au changement de la connaissance et de la mentalité des fonctionnaires et des journalistes s'appuie sur une « progression automatique » entre la constatation d'une situation injuste et l'intervention afin d'y mettre fin;
- 2) Les résultats empiriques montrent qu'il est injustifié de mettre l'accent sur une nouvelle allocation de pouvoirs comme facteur de changement.

## ■ LES FONCTIONS ET LE CADRE INSTITUTIONNEL DE L'ÉVALUATION

La Banque mondiale met systématiquement en œuvre des évaluations internes ex post pour les programmes de développement. Dans le schéma institutionnel, la fonction évaluative a été distinguée des fonctions budgétaires et d'audit. En 1970, Robert McNamara, président de la Banque mondiale, a créé un groupe autonome d'évaluation à l'intérieur du Département de la planification et du budget. La première unité d'évaluation de la Banque a été mise sur pied à l'intérieur du Département de l'évaluation des opérations de la Banque (OED).

En 1975, à la demande des actionnaires, l'OED est devenu indépendant du Conseil des directeurs exécutifs de la Banque mondiale (Picciotto, 1997). Actuellement, outre l'OED, deux autres organisations de taille, de champs de compétence et d'importance plus modestes ont, à l'intérieur de la Banque mondiale, la possibilité d'effectuer des évaluations ex post. Le Groupe d'évaluation des opérations de la Société Financière Internationale et l'unité d'évaluation de l'Institut de la Banque mondiale (WBI) évaluent respectivement les investissements privés et la formation et les séminaires délivrés dans les pays en voie de développement.

La Banque mondiale étant une organisation « basée sur la connaissance », où la gestion est menée à partir de l'information et du savoir (Kapur, Webb, Lewis, 1997), l'évaluation ex post s'ajoute à la somme, à la fois considérable et très diversifiée, des études techniques, chevauchant ainsi différents départements d'analyse de politiques, de recherche et d'évaluations ex ante. Par exemple, le Groupe d'assurance qualité (QAG) effectue des contrôles qualités ex ante à la conception de tous les documents de la Banque ; le Groupe de Politiques et Stratégies Opérationnelles (OPS) est responsable de la formulation de la politique, de son examen, et de la conformité au suivi ; le WBI et la Commission européenne fournissent respectivement des formations et des recherches dans l'économie du développement<sup>1</sup>.

Dans ce contexte, l'OED jouit d'un statut particulier. Son indépendance est basée sur le fait que, au niveau organisationnel, la fonction d'évaluation dépend directement du Conseil avec un accès illimité à toutes les données et à tous les employés de l'organisation. Les évaluations de l'OED sont présentées et commentées dans le cadre du sous-comité du Conseil pour l'efficacité en matière de développement (CODE) et sont ensuite débattues lors de réunions du Conseil.

Le mandat officiel établit que le directeur général de l'OED « relève directement des directeurs exécutifs de la Banque » ; il est nommé par le Conseil pour un mandat renouvelable de cinq ans et ne peut être démis de ses fonctions qu'à la suite d'une décision du Conseil. Par la suite, il lui est impossible d'occuper un poste à la Banque, même s'il s'agit d'une position qu'il détenait avant sa nomination (OCDE, 1996).

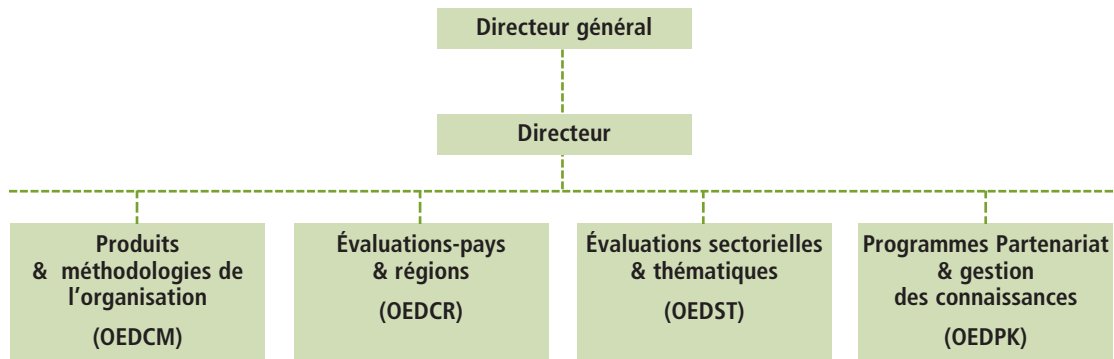
Contrairement aux autres unités et départements mentionnés plus haut qui luttent pour obtenir une part d'un budget global en très forte baisse, les ressources financières de l'OED sont allouées directement par le Conseil en fonction du programme de travail. L'engagement institutionnel en faveur de l'évaluation conduit le président à autoriser cette indépendance financière, libérant ainsi l'OED de la compétition budgétaire interne. Cette organisation institutionnelle permet également de qualifier l'action de l'OED de « fonction évaluative externe » en regard des principes généraux de contrôle des administrations publiques (*Generally Accepted Government auditing Standards*) émis par le *General Accounting Office* des États-Unis. Ceci renforce la crédibilité de l'OED comme l'une des meilleures organisations évaluatives dans le domaine de l'aide au développement (Chelimsky et Shadish, 1997<sup>2</sup>).

Le personnel de l'OED est composé d'environ 70 personnes, réparties en trois divisions et une unité se consacrant aux fonctions supports, comme le montre le schéma.

<sup>1</sup> Observations et analyse documentaire.

<sup>2</sup> Le Comité d'aide au développement de l'OCDE a également publié « *Principles for Evaluation of Development Assistance* » qui affirme que l'impartialité et l'indépendance sont mieux servies lorsque la fonction évaluative est clairement séparée des organes chargés de planifier et de gérer l'aide au développement. Il est également spécifié que cette séparation peut être assurée au moyen d'un service d'évaluation centralisé placé sous la responsabilité directe du Conseil des directeurs ou du Conseil des gouverneurs de l'institution.

## SCHÉMA D'ORGANISATION DU DÉPARTEMENT DE L'ÉVALUATION DES OPÉRATIONS DE LA BANQUE MONDIALE (OED)



Du point de vue fonctionnel, le Département remplit deux tâches principales. En premier lieu, le système de développement de l'évaluation est pris en charge par le service Produits et méthodologies de l'organisation (OEDCM) et les programmes Partenariat et gestion des connaissances (OEDPK). Il s'agit de développer les capacités d'évaluation et de renforcer l'acceptation des recommandations de l'évaluation grâce à des stratégies de publication et de diffusion. En second lieu, la production d'études est effectuée par les équipes d'Évaluations-pays et régions (OEDCR) et des Évaluations sectorielles et thématiques (OEDST). L'OED est utilisé en tant qu'aide à la décision concernant les fonds que la Banque prête (25 à 30 milliards de dollars chaque année) ou dépense en administration et en services non liés aux prêts (environ 1,5 milliard de dollars US chaque année) (Piccioto, 1997).

L'OED est mandaté pour examiner une très grande diversité d'activités de développement. Les prêts de la Banque mondiale sont utilisés pour exécuter des projets de reconstruction et de développement dans une large variété de secteurs et pour améliorer les plans d'action et les programmes dans des cas très divers d'économie en développement ou en transition. L'institutionnalisation de la fonction évaluative à l'intérieur de la Banque mondiale trouve son origine dans le besoin d'assurer un système de contrôle financier formel, mais aussi dans celui de satisfaire aux demandes de reddition de comptes

de la part des États contributeurs, comme l'écrit Piccioto (2002), ancien directeur général de l'OED :

**Rendre des comptes au sujet d'une assistance au développement réactive, efficace et axée sur les résultats est devenu un impératif fondamental pour la Banque mondiale et ses partenaires dans le développement. L'efficacité, la transparence et des résultats documentés sont de première importance. Une fonction évaluative rigoureuse, pertinente et indépendante est devenue essentielle à la crédibilité de l'institution (Picciotto, Wiesner, 1998).**

En 1999, l'OED a mené des inspections et des audits de performance pour 290 opérations de développement déjà achevées. Ces opérations couvrent des secteurs majeurs de l'économie dont l'agriculture, l'énergie, les finances, le transport, la distribution d'eau et le développement urbain. Elles sont mises en œuvre par l'intermédiaire d'une vaste gamme d'instruments : investissements dans des projets spécifiques, prêts à un secteur ou à une économie, schémas d'intermédiation financière, ajustements de politiques, prêts d'urgence. L'OED note la performance de chaque projet selon le niveau de décaissement des crédits prévus.

La base de données résultant de ces évaluations contient actuellement le schéma d'ensemble, les caractéristiques de performance et les leçons à tirer de l'expérience pour à peu près 4800 projets de développement (Picciotto, 2002, commentaires).

Un quart des projets achevés sont évalués selon un «audit de performance». Contrairement à ce que la terminologie employée pourrait faire croire, ces rapports comprennent l'évaluation de l'efficacité du projet, son efficacité, sa pertinence, sa viabilité et son impact en termes de développement institutionnel.

Afin de préparer ces audits de performance, les évaluateurs de l'OED étudient les documents afférents au projet, mènent des entrevues avec le personnel opérationnel et, dans la majorité des cas, effectuent des vérifications de terrain dans les pays emprunteurs. Le personnel concerné, les emprunteurs et les partenaires reçoivent une version préliminaire de l'audit. Les commentaires des emprunteurs et des prêteurs associés sont incorporés au document qui est envoyé au Conseil de la Banque mondiale.

Lorsque le directeur général décide de diffuser un rapport, il est largement distribué à l'intérieur de la Banque mondiale ainsi qu'aux autorités concernées dans les pays membres. Comme le déclare Picciotto: «Au moment de sélectionner les projets faisant l'objet d'un audit de performance, la priorité est donnée aux projets innovants, de grande ampleur ou complexes; à ceux pour lesquels les directeurs exécutifs demandent un audit; ou à ceux qui, individuellement ou du fait de leur appartenance à un ensemble de projets, sont susceptibles de fournir des enseignements importants. Les projets, les sujets et les approches analytiques des programmes de travail des audits de performance sont choisis afin de fournir des données utiles pour les autres études évaluatives.» (Picciotto, 2002).

Ces analyses servent à mesurer et à préconiser des changements dans les activités de l'organisation ainsi qu'à suivre, dans le temps et de manière globale, comment les projets sont mis en œuvre en différents endroits. Elles servent aussi à faire ressortir l'opinion des bénéficiaires à l'égard d'une unité opérationnelle et de ses projets et d'évaluer les dépenses en fonction de la planification établie. Au-delà de cette activité d'évaluation des projets, l'OED consacre une part constante de ses moyens à des études recoupant différents types de données, incluant ainsi des évaluations de procédures et d'incidences, des estimations de programmes-pays ainsi

que des révisions de politiques globales ou régionales, comme c'est le cas pour l'étude de foresterie.

Le tableau 3 résume les résultats, l'optique, le but, la méthodologie et les ressources utilisées pour les évaluations de l'OED. L'approche et la méthodologie varient selon l'optique de l'évaluation et l'objet étudié. Alors que les approches participatives ont été de plus en plus utilisées dans la Banque en association avec des méthodes d'enquêtes qualitatives – par exemple, entretien en profondeur, concentration sur un groupe, étude de terrain et étude de cas –, la perspective de l'évaluation va de l'estimation d'impact, de calculs coûts-bénéfices et coûts-efficacité et de la pertinence de stratégies d'ensemble et de politiques thématiques à des projets d'audits de performance et d'examens de mise en œuvre. Les évaluations basées sur des données de seconde main, comme les évaluations de projets hors terrain, les documents de programmes et les analyses statistiques de bases de données compilant des évaluations de projets viennent en complément d'études de terrain et d'autres analyses empiriques, comme l'étude de foresterie et l'analyse des impacts des Revues de dépenses publiques.

Les études thématiques et les révisions de politiques cherchent à développer les indicateurs et les objectifs de performance nécessaires à l'amélioration de l'efficacité et de la réceptivité institutionnelle. Le but est ici de déterminer comment la coopération survient ou ne survient pas, entre les organismes donateurs collaborant à un programme. Il s'agit également de déterminer les systèmes de données nécessaires pour comprendre ce qui se passe dans un programme ou une politique.

Les ressources humaines de l'évaluation sont principalement internes ou agrégées ponctuellement par contrat, particulièrement pour les études hautement qualifiées ou de grande ampleur. Les inspections externes sont confiées à des universitaires ou à des experts. Néanmoins, on assiste de plus en plus fréquemment à la création d'un comité consultatif indépendant qui garantit l'impartialité de jugement et une expertise technique hautement reconnue<sup>3</sup>. Enfin, comme le montre le tableau 3, l'optique d'évaluation de l'OED concilie l'attention portée à la reddition de comptes strictement financière avec

**TABLEAU 3 : DIFFÉRENTES APPROCHES DE L'ÉVALUATION ET ORGANISATIONS RESPONSABLES**

TYPE D'ÉVALUATION	OPTIQUE D'ÉVALUATION	BUT DE L'ÉVALUATION	MÉTHODE ÉVALUATIVE	EXPERTISE DE L'ÉVALUATION	ORGANISATION ÉVALUATRICE
Évaluation sommaire	Reddition de comptes sur le projet financier	Optimisation des ressources Développement des connaissances	Analyse documentaire	Économistes/ Auditeurs/ Ancien personnel opérationnel de la Banque	OED interne
Audit de performance	Reddition de comptes et performance du projet/ Rentabilité économique	Optimisation des ressources Développement des connaissances	Analyse coût-bénéfice <i>ex post</i> / Analyse de performance /Étude de terrain	Économistes/ Gestionnaires expérimentés de la Banque	OED interne
Évaluation – pays	Reddition de comptes institutionnelle/ Rentabilité économique de l'emprunteur	Développement/ Amélioration de la programmation des prêts de la Banque	Étude de terrain/ Analyse de performance de la Banque/ Analyse de performance de l'emprunteur/ Indicateurs de performance	Économistes	OED interne/ Consultants externes ou chercheurs pour la collecte de données sur le terrain
Étude thématique	Reddition de comptes institutionnelle / Développement	Renforcement des capacités de planification de la politique sectorielle	Étude de terrain/ Analyse de performance/ Approches participatives et consultatives/ Révision par les pairs	Économistes/ Spécialistes des sciences sociales	OED interne/ Sous-traitance pour les études de cas pays/ Comité consultatif
Examen de politique	Développement	Renforcement des capacités de planification et d'analyse	Étude de terrain/ Analyse de performance/ Approches participatives et consultatives/ Révision par les pairs	Économistes/ Spécialistes des sciences sociales / Chercheurs/ Personnel de support de la Banque/ Universitaires	OED interne/ Sous-traitance pour les études de cas pays / Comité consultatif

<sup>3</sup> Dans le cas de l'étude de foresterie de l'OED et de l'évaluation des Revues de dépenses publiques (PER), les évaluateurs principaux ont mis en place un comité consultatif afin de bénéficier d'une expertise supplémentaire et d'un avis indépendant. L'analyse de certains PER a été spécifiquement confiée à des groupes d'évaluateurs indépendants y compris des universitaires reconnus en finances publiques. Il leur a été demandé de suivre strictement la méthodologie qui leur a été fournie. Les rapports d'évaluation ont été distribués aux gestionnaires de projet de la Banque afin de recueillir leurs commentaires. Comme le souligne l'évaluation, cette méthode a permis de renforcer l'équipe évaluatrice en équilibrant l'expertise technique des conseillers externes avec l'expérience de la Banque des évaluateurs de l'OED et les connaissances des fonctionnaires des pays en voie de développement. L'examen du portefeuille des projets de foresterie a été fait en collaboration avec un comité consultatif. Les quatre membres de ce comité ont conseillé l'OED sur l'avant projet, le choix des pays pour les études de cas, les découvertes préliminaires à partir du portefeuille de projets, le plan de concertation global de l'étude et le rapport final. Les membres du comité ont été en contact régulier avec l'équipe de l'OED. Dans les deux cas, le comité consultatif a apporté une valeur ajoutée au niveau méthodologique dans le processus d'évaluation et de la légitimité des conclusions.

le besoin de promouvoir le développement de la connaissance, l'apprentissage et le renforcement institutionnel<sup>4</sup>.

Dans ce contexte, si le modèle institutionnel, les fonctions organisationnelles et les schémas spécifiques des études évaluatives nous permettent de distinguer le département chargé de l'évaluation des autres départements de la Banque, ceci ne garantit pas sa légitimité interne ou son « avantage compétitif » vis-à-vis d'autres services ayant un mandat similaire. En conséquence, au-delà de cette indépendance de l'évaluation, nous ferons porter notre attention sur sa légitimité institutionnelle et politique réelle à l'intérieur de la Banque. Notre objectif est de comprendre comment et dans quelles circonstances l'évaluation joue un rôle d'agent du changement organisationnel, renforçant ainsi son influence sur le processus interne de prise de décision.

### ■ L'ÉVALUATION, ENTRE TECHNOCRATIE ET GOUVERNANCE DÉMOCRATIQUE

Dans le droit fil des recherches les plus récentes en matière de comportement organisationnel, nous devons considérer l'OED comme enchâssé dans l'environnement global de la Banque mondiale afin de saisir – pour paraphraser Powell et DiMaggio (1991) – la pression isomorphe modelant ses processus organisationnels de travail et ses productions. L'isomorphisme est « un processus contraignant qui force une unité dans une population à ressembler aux autres unités confrontées au même ensemble de conditions environnementales » (Powell et DiMaggio). Ce concept nous aide à expliquer l'homogénéité de ces formes et pratiques organisationnelles plutôt que leurs variations. Powell et DiMaggio postulent que des organisations à l'origine très différentes sont, une fois qu'elles se trouvent dans le même secteur d'activité, structurées dans un champ d'actions communes, défini implicitement ou codifié explici-

tement par les règles professionnelles ou organisationnelles. L'homogénéisation de leurs modes de travail et de leur production se fait à travers des relations d'autorité ou professionnelles ou les deux.

Les équipes des départements de la Banque mondiale spécialisées dans l'analyse de politiques, la recherche et l'évaluation démontrent un certain nombre de similarités entre elles. En premier lieu, l'« isomorphisme coercitif » résulte du schéma organisationnel global et du cadre légal de la Banque mondiale. Ces éléments influent sur le travail de l'OED par l'intermédiaire de procédures, de règles et de structures dont les variations limitées entre les différents services et unités sont compatibles avec les politiques internes établies et légitimées.

Les pratiques comptables, les mesures et l'évaluation de la performance, les politiques en terme de ressources humaines et les prévisions budgétaires sont imposées à l'OED par l'autorité directe des gestionnaires de haut niveau et du Président. En second lieu, l'« isomorphisme compétitif » provient de l'« environnement technique basé sur la connaissance » de la Banque mondiale qui façonne l'organisation interne des équipes en fonction des attentes normatives et culturelles (Powell et DiMaggio, 1991). Les fonctions professionnelles et les productions demandées définissent les conditions et les méthodes du travail d'équipe, y compris lorsqu'il s'agit d'évaluation. Par exemple, les processus consultatifs et participatifs, les recherches évaluatives, les analyses formelles basées sur des statistiques ou des faits empiriques sont typiques des méthodes de travail de l'OED comme du WBI, du QAG, de l'OPS et du DEC<sup>5</sup>.

Du point de vue des personnes que nous avons interrogé, les rapports d'évaluation sont de même nature que les notes de service, les documents stratégiques, les analyses économiques et les recueils statistiques. Toutes ces productions des différents services de la Banque mondiale sont en concu-

<sup>4</sup> Entrevues semi-directives avec des évaluateurs de l'OED.

<sup>5</sup> Par exemple, les Stratégies d'assistance au pays (CAS), les Revues de dépenses publiques (PER), les Évaluations de la politique et des institutions nationales (CPIA), les Documents sectoriels de lutte contre la pauvreté (PRSP), les Cadres de développement intégrés (CDF) ne représentent qu'une partie des documents analytiques officiels produits par la Banque et les pays emprunteurs en prévision de prêts.



rence pour attirer l'attention et exercer une influence sur la politique menée, dans un contexte où les données, les idées et les points de vue sur des modes d'actions alternatifs se multiplient dans l'organisation. Par exemple, l'étude de foresterie de l'OED coexiste avec au moins 17 autres analyses réalisées sur le même thème dans le cadre du processus d'inspection de la politique forestière, initiée en 1998 par le département de l'environnement<sup>6</sup>. L'évaluation des PER, complétée par l'OED en 2002, a été mise à jour et intégrée par l'OED pour l'analyse de qualité *ex ante* en 2000.

En accord avec les conclusions de Nonaka (1994) sur la création du savoir organisationnel et de Kapur et al. (1997) qui ont décrit les caractéristiques du management par la connaissance à la Banque mondiale, nous pouvons conclure que ce sont les experts ayant le plus d'information et de connaissances qui prennent l'initiative d'agir. La professionnalisation est largement répandue et explique l'homogénéité des tâches accomplies par le personnel de la Banque. Cette homogénéité résulte d'une base cognitive partagée par tous qui légitime leur autonomie et leur satisfaction à l'intérieur de la Banque mondiale. L'éducation classique – principalement en économie<sup>7</sup> – et l'interaction continue avec les universités les plus prestigieuses<sup>8</sup> expliquent l'existence d'une connaissance professionnelle et technique donnée, qui renforce l'« iso-

morphisme compétitif » interne. « Plus une organisation se base sur les diplômes universitaires pour recruter son personnel, plus elle a de chances de devenir comme les autres organisations du même secteur d'activité. » écrivent Powell et DiMaggio (1991).

Ce qui fait que l'OED est en concurrence et ressemble à toutes les autres équipes de la Banque, c'est que le profil professionnel des évaluateurs renvoie aux qualifications et aux compétences que l'on retrouve ailleurs dans la Banque. Les évaluateurs sont des économistes ayant une expérience dans la gestion de projets de développement et/ou dans la recherche et l'analyse de politiques. Les gestionnaires de différents secteurs et aires géographiques de la Banque mondiale joignent souvent l'OED à la fin de leur carrière, même si la politique récente de l'OED entend promouvoir l'échange d'expériences par la notion de « tourniquet ». En effet, les évaluateurs doivent régulièrement aller et venir entre l'OED et les unités opérationnelles afin d'être en mesure d'évaluer des opérations de développement complexes. Selon les gestionnaires de programmes, une telle rotation peut enrichir l'expérience professionnelle des évaluateurs, mais elle est également susceptible d'affaiblir l'indépendance des études menées du fait de conflits d'intérêts. C'est la raison pour laquelle la procédure d'allocation du travail de l'OED prévoit d'éviter aux évaluateurs

<sup>6</sup> Le processus consultatif, qui a été mené par l'équipe d'évaluation tant à l'intérieur qu'à l'extérieur de la Banque dans six pays en voie de développement, chevauchait d'autres types de consultations effectuées par des réunions officielles, des discussions par Internet et des discussions informelles au service central du Département du développement socialement et écologiquement durable de la Banque mondiale (ESSD).

<sup>7</sup> Par exemple, la plupart des répondants de la Banque mondiale sont titulaires d'une maîtrise en sciences sociales ou l'équivalent d'une université nord-américaine ou de l'Europe de l'Ouest. Dans des départements comme l'OED, le DEC ou le WBI, plus spécialisés que le reste de l'organisation, les titulaires de doctorat sont fréquents.

<sup>8</sup> Les liens avec les universités canadiennes et américaines par l'intermédiaire des associations des étudiants diplômés sont également fréquents, ces organisations pouvant également servir de lieu de rencontre permettant d'obtenir un emploi. Les relations avec les universités vont de la simple diffusion ou présentation d'articles scientifiques à l'heure du dîner à la mise sur pied de comités consultatifs composés de professeurs reconnus dans le domaine concerné.

d'expertiser des opérations dans lesquelles ils ont directement participé avant d'être au service du Département.<sup>9</sup>

Toutefois, les entrevues effectuées révèlent que la réputation de l'OED à l'intérieur de la Banque mondiale est ambivalente, ce qui nous fournit une indication ambiguë sur le niveau de soutien culturel envers le rôle de l'évaluation et le département qui la conduit. D'un côté, le service d'évaluation jouit d'une grande autonomie institutionnelle, technique et financière. D'un autre côté, les employés de la Banque sont divisés sur la perception du statut et du prestige attachés, dans une perspective d'avancement de carrière, à la profession d'évaluateur. Les déclarations suivantes reflètent bien ce point de vue :

**Les évaluateurs de l'OED sont comme des historiens (...) dans une armée; est-ce que vous mettriez un historien à la tête de vos hommes...? Il devrait y avoir un système de rotation pour faire en sorte que les personnes de l'OED soient obligées après un certain temps de retourner aux opérations sans qu'elles en pâtissent.**

D'une manière générale, l'OED est un bon endroit pour un employé de la Banque qui ne veut pas trop voyager ou qui veut prendre sa retraite prochainement.

Les gestionnaires de programmes interviewés ont minimisé le prestige associé aux évaluations, en les opposant à la gestion directe de prêts. Les experts hautement spécialisés – comme les ingénieurs environnementaux qui œuvrent à la construction de grandes infrastructures ou de barrages ou les économistes travaillant sur la structure des dépenses des pays en voie de développement – ont discrédité

la procédure d'évaluation en raison du manque d'analyse technique. Le mode d'évaluation leur apparaît insuffisant et inadéquat comparativement aux analyses économiques, techniques et financières entreprises généralement par la Banque. Comme le dénoncent certaines personnes interrogées, les évaluations présentent souvent de l'information déjà connue :

**Le rapport (PER) émet des généralités. Bien qu'il résume la situation et qu'un résumé soit toujours utile, il est rempli de sens commun.**

**Nous savons comment construire un barrage. Nous connaissons la problématique du déplacement des populations. Je suis un professionnel ayant une excellente connaissance relative à la construction des barrages, car je travaille sur cette problématique depuis 12 ans.**

Malgré cela, tous les répondants ont reconnu la valeur et l'importance de la fonction d'évaluation de la Banque. Le mandat de l'OED est perçu par tous comme nécessaire, car il permet un juste équilibre des pouvoirs et une bonne gouvernance démocratique.

**L'évaluation permet la mise en place de mesures de contrôle, lesquelles contrebalancent la prise de décision de gestion.**

**L'évaluation doit permettre d'instaurer dans la Banque un équilibre démocratique entre les gestionnaires et les employés sur le terrain.**

Contrairement à la croyance traditionnelle qui veut que l'évaluation interne ne remette pas fondamentalement en question les politiques organisationnelles, se pliant ainsi aux demandes des gestionnaires

<sup>9</sup> Entrevues avec les évaluateurs et la direction du département d'évaluation. Le directeur général affirme : « La solution pour éviter les conflits d'intérêts réside dans l'application des recommandations de l'IIA qui stipule que "les auditeurs internes doivent s'abstenir d'évaluer les opérations spécifiques dont ils étaient précédemment responsables." De plus, afin d'éviter ce risque, sur les deux dernières années fiscales (2001 et 2002), environ un tiers des personnes embauchées à l'OED venaient de l'extérieur de la Banque, et les quatre cinquièmes des personnes qui quittent le département quittent aussi la Banque mondiale. Un mélange entre recrues internes et externes est considéré comme étant le meilleur moyen de mener des évaluations de grande qualité. C'est pourquoi le nouveau bureau d'évaluation du FMI a été pourvu de personnes de l'extérieur et de l'intérieur de l'organisation. Michael Scriven, un des meilleurs experts en matière d'évaluation, souligne les avantages d'un recrutement interne : « les individus concernés ont tendance à savoir bien mieux naviguer dans l'organisation que les personnes venant de l'extérieur et évitent ainsi des erreurs dues à une méconnaissance de l'environnement. »

(Boyle, Lemaire, 1999), la plupart des entrevues soutiennent l'idée que l'indépendance est le facteur clef. Selon la majorité des répondants, c'est l'indépendance institutionnelle de l'OED qui lui confère de la légitimité.

Nombre de facteurs mènent à cette indépendance. Tout d'abord, le cadre institutionnel de l'évaluation fait qu'elle ne dépend pas d'interventions du président ou de décideurs haut placés. Selon son mandat institutionnel, l'OED est responsable devant le conseil d'administration de la Banque, un organe suffisamment éloigné des activités opérationnelles. Cette situation permet de réduire le biais inhérent lors de la mise en place d'un processus de révision des programmes. Ensuite, l'OED fait montre d'indépendance (1) en préparant son programme et son agenda de travail; (2) en choisissant les projets qu'il souhaite évaluer; (3) en subventionnant les études ainsi qu'en choisissant leurs contraintes de délais; (4) en faisant le choix de la méthodologie appropriée pour la collecte des données; (5) en rapportant des résultats à travers différents moyens de transmission des données. Trois des quatre études de cas, soit celles de l'industrie forestière, des PER et de la construction de barrages, ont permis de démontrer l'indépendance de l'OED. Les évaluateurs ont fait montre d'indépendance en créant le cadre d'évaluation, en formulant la question de recherche principale, en choisissant la méthodologie adéquate, en effectuant l'analyse des données, en rédigeant et en diffusant le rapport final.<sup>10</sup>

Enfin, les entrevues ont suggéré que le leadership démontré par le département d'évaluation joue un rôle non négligeable du point de vue de la politique, en questionnant la relation entre les politiques et les objectifs désirés par la Banque et en assurant une crédibilité, une intégrité et une neutralité dans les jugements. Durant les réunions du conseil d'administration, le directeur général joue le rôle d'une tierce partie qui n'hésite pas à critiquer les décisions de gestion, à questionner les prémisses implicites à la base des programmes de développement et, finalement, à arbitrer entre les priorités du conseil

et celles relatives aux besoins des gestionnaires.

Par exemple, le 14 février 2001, lors d'une réunion du conseil où devaient se discuter les constats et les recommandations de la Commission mondiale sur les barrages, le directeur général a rappelé à tous les participants les résultats d'une analyse coûts-bénéfices réalisée en 1997. Il a défendu la méthode utilisée par l'OED, laquelle prend en considération les coûts et les bénéfices ainsi que les implications sociales, environnementales et économiques associées à la construction des barrages. Selon lui, cette méthode est valable pour l'évaluation effective des barrages. Par son intervention, il a réussi à harmoniser le dialogue entre, d'une part, l'opposition des gestionnaires aux recommandations faites par la Commission mondiale des barrages et, d'autre part, l'intention du conseil de garder le débat ouvert sur les sources d'énergie alternatives à l'énergie hydro-électrique.

Le leadership est donc un facteur essentiel pouvant donner de la légitimité politique à l'évaluation. Sur des enjeux critiques de développement, le directeur joue donc un rôle « d'avocat du diable » vis-à-vis des gestionnaires de haut niveau et des décideurs. Comme il est indiqué dans la citation ci-dessous, alors que la légitimité politique s'acquiert au sommet de l'organisation et aux niveaux les plus hauts de la prise de décision, elle fait défaut aux niveaux inférieurs de gestion.

**Le leadership démontré par l'OED contrebalance la plupart des erreurs faites dans les rapports d'évaluation.**

Étant donné que l'évaluation est mal reçue, que plusieurs craignent ses résultats et doutent que ses recommandations soient suffisamment utiles pour entraîner un réajustement des programmes, les entrevues ont laissé entendre que, bien souvent, l'interaction entre les évaluateurs et les gestionnaires de programmes est limitée. Trois interviewés ont noté que les évaluateurs faisaient preuve systématiquement d'un penchant hostile aux gestionnaires de programmes.

<sup>10</sup> Entrevues semi-structurées.

Parfois les évaluateurs de l'OED étaient arrogants et agressifs et semblaient soutenir qu'eux seuls connaissaient la vérité.

Selon les gestionnaires de programmes, la consultation et la participation au cadre d'analyse d'évaluation sont parfois insuffisantes pour créer des conditions propres à la confiance et à l'échange d'information. Il ne suffit pas en effet de consulter les gestionnaires ou d'utiliser la révision par les pairs. Il conviendrait plutôt de s'assurer que leur participation soit effective à toutes les étapes du processus d'évaluation : (1) dès le début de l'évaluation, afin que tous s'accordent sur les objectifs de l'étude et les questions de recherche pertinentes ; (2) lorsque le projet est en cours d'évaluation afin que les commentaires reçus puissent servir à une rétroaction efficace sur le travail effectué ; (3) à la fin du processus d'évaluation pour que les gestionnaires puissent partager leurs commentaires sur le processus.<sup>11</sup>

Certains interviewés furent vexés devant l'impossibilité de discuter les conclusions et les résultats du rapport avec lesquels ils étaient en désaccord. Comme le démontre le cas de l'industrie forestière, l'évaluation participative a porté ses fruits en créant les conditions d'une crédibilité et d'une légitimité politique plus importantes. Ceci eut pour effet de faciliter le partage et l'utilisation de la connaissance évaluative utile à la prise de décision stratégique à tous les niveaux de la gestion. Le manque d'interaction entre les gestionnaires de programmes et les évaluateurs donne lieu à ce que certains auteurs nomment l'information asymétrique, créant ainsi un « dilemme vertical » (Knott, Miller, 1998). Comme Knott et Miller semblent le démontrer, la hiérarchie et la délégation peuvent mener à une action organisationnelle incohérente. Afin de mieux comprendre cet aspect de l'évaluation, il convient de s'intéresser maintenant à l'analyse des barrières institutionnelles, culturelles et politiques responsables du manque de lien organisationnel (Goodman, 2000).

## ■ LES BARRIÈRES CULTURELLES À L'UTILISATION DE L'ÉVALUATION

Selon la plupart des personnes consultées, les évaluations restent pertinentes et utiles pour les gestionnaires de programmes et les employés de la Banque mondiale responsables de la révision des politiques. Les réponses indiquent qu'il existe une forte corrélation entre la contribution d'une étude à la gestion d'un programme et la probabilité d'utiliser cette étude. Étonnamment, les évaluateurs et les gestionnaires de programmes s'accordent sur le « public cible » auquel devraient s'adresser les rapports d'évaluation. Bien que les évaluateurs soient conscients que le mandat institutionnel de l'évaluation est de rapporter l'information au niveau le plus haut de la prise de décision, tous s'entendent pour conclure que c'est sur le plan de la gestion que l'évaluation trouve sa plus grande utilité. Extraits de quelques réactions :

**Les vice-présidents, les directeurs-pays, les réseaux, de manière à partager l'expérience passée de la Banque...**

**Un mois m'apparaît être un siècle, mais je me souviens de la discussion sur les barrages (...) Je suis un expert de la puissance énergétique.**

**Les nouveaux employés se doivent d'étudier l'information de l'OED afin de se faire une idée du travail de la Banque. (...) L'information de l'OED sera utile aux nouveaux gestionnaires.**

Il ressort des entrevues que les gestionnaires utilisent les évaluations afin d'encadrer le processus décisionnel et le plan de gestion. Pour les gestionnaires chevronnés (directeurs-pays, chefs de secteur et membres du conseil), les évaluations de l'OED sont perçues de manière instrumentale, comme un mécanisme de rétroaction permettant de guider la prise de décision. Selon eux, un compte rendu de l'expérience passée est particulièrement utile pour permettre l'amélioration de la politique, comme cela a été démontré dans le cas de l'étude sur l'industrie forestière. Pour que la perception de l'évaluation soit maximale chez les répondants, il importe d'y inclure des recommandations pour la mise en place d'actions spécifiques.

<sup>11</sup> Entrevues avec les gestionnaires de programmes.

Dans trois des quatre études de cas, les répondants ont mentionné que les recommandations qui suggéraient des actions concrètes risquaient fort d'être prises en compte. Il importe cependant de s'interroger sur la valeur des recommandations comme véritablement incitatives au changement. Pour les responsables en charge des projets ou de la gestion du programme, il convient d'agir selon les recommandations. À la question de savoir si les recommandations offrent une ligne directrice qui puisse être suivie, la majorité des gestionnaires de programmes ont répondu comme suit :

**Nous les mettons déjà en place (...) On se doit de rendre les PER plus participatives et de les centrer sur certains secteurs spécifiques (...) C'est la manière dont les PER sont faites !**

Selon ce qui ressort de certaines entrevues, les gestionnaires expérimentés perçoivent la ligne directrice de l'évaluation comme étant plus utile pour leurs subordonnés que pour eux-mêmes.

**Je prends pour acquis que mes employés intégreront les leçons apprises dans la formulation des nouveaux projets.**

**Les gens en charge d'écrire la nouvelle stratégie forestière se doivent de prendre en considération les recommandations de l'OED.**

Suivant la thèse de Weiss et Bucuvalas (1980a) sur le comportement des fonctionnaires, il semblerait que ceux-ci accordent peu d'importance à l'évaluation lorsqu'elle les concerne, mais qu'ils y portent un plus grand intérêt lorsqu'elle concerne le comportement des autres employés. En rapport avec la thèse sur la conformité bureaucratique, les gestionnaires de haut niveau ont paru faire preuve d'inertie et manquer de réponses stratégiques face aux suggestions des évaluations. Cette réaction est compatible avec les systèmes de connaissances propres aux organisations bureaucratiques comme, par exemple, la croyance en des sphères de compétence délimitées, la hiérarchie, l'importance d'une procédure propre et l'impersonnalité.

Les organisations bureaucratiques prennent pour acquis que les individus sont des êtres abstraits, lesquels sont déterminés par la position qu'ils occupent (Berger et Luckmann, 1968). Berger et Luckmann insistent sur les systèmes cognitifs partagés, conçus

comme des structures objectives et externes qui définissent la réalité organisationnelle. La culture n'est pas confinée dans le subjectif, les pensées profondes de l'individu ou une vague notion de conscience collective, mais elle est plutôt reconnue comme représentant une réalité objective propre. Plus les catégories et les systèmes de croyances sont institutionnalisés et plus les actions humaines seront définies à travers des routines tenues pour acquises (Berger et Luckmann, 1968). La bureaucratie, et la forme qu'elle prend, devient donc une barrière institutionnelle affectant l'organisation, les systèmes de croyances, les attitudes mentales et la culture propre à ses membres. Ceci aura pour effet de réduire la capacité de l'information évaluative à être prise en compte lors des décisions de gestion.

Parallèlement, il semblerait que la culture propre à la Banque mondiale influence directement les études évaluatives, et ce, en renforçant la distance psychologique déjà démontrée jusqu'ici. Étant donné que les évaluations sont axées sur certaines valeurs et que l'on reproche aux évaluateurs d'être biaisés, les gestionnaires de programmes et les décideurs regrettent le mode d'action normatif prescrit par les évaluations, qui ne s'appuierait pas, d'après eux, sur des preuves empiriques. Selon certains, les recommandations ne doivent être comprises que comme de « bonnes intentions ».

De manière plus spécifique, les gestionnaires de programmes se sont intéressés à la détermination du programme des évaluations, ainsi qu'aux préférences des évaluateurs et du département d'évaluation. À l'opposé, le département interne d'évaluation estime accepter les valeurs de la Banque et les paradigmes dominants relatifs au développement sans en questionner la validité scientifique et empirique, manquant ainsi d'analyse critique et de distance. Les valeurs et les visions du monde qui sont propres à la communauté du développement international influencent grandement la culture entourant le processus d'évaluation, affectant ainsi la crédibilité de ce processus.

**De plus en plus, l'OED produit des études ayant des sujets de développement à la mode, tels que les CDF ou les autochtones. Il fait ensuite des recommandations ayant pour but d'augmenter la participation, de promouvoir l'égalité des sexes et**

**bien d'autres sujets à la mode en développement (...)  
Cependant, dans dix ans, on se rendra compte que la participation n'est pas suffisante (...)  
L'OED se doit d'être moins influencé par les sujets en vogue dans ses recommandations.**

Une fois de plus, on réalise à quel point la culture influence la vie organisationnelle, et ce, indépendamment de la structure organisationnelle, de la disponibilité des ressources et des exigences techniques. La culture doit être vue ici comme « une boîte à outils pleine de symboles, d'histoires, de rituels et de visions du monde que les organisations peuvent utiliser de différentes manières pour régler une variété de problèmes » (Meyer et Rowan, 1991). Par exemple, on peut penser aux dommages environnementaux liés à la construction de barrages ou aux arbitrages quant à l'allocation des dépenses publiques. Ainsi donc, le système de croyances et l'ensemble des connaissances techniques propres à la Banque mondiale deviennent des facteurs institutionnels qui orientent et modèlent l'ensemble du processus d'évaluation.

Qui plus est, le département d'évaluation participe au système cognitif et normatif de la Banque et représente la culture et les valeurs propres à cette organisation internationale. Néanmoins, le fait d'adhérer à des connaissances et à des valeurs organisationnelles ne représente pas un problème en soi. Si on considère que la connaissance est une construction sociale, il apparaît que les valeurs, les opinions, les croyances, les préférences, les intérêts, les désirs et les besoins ont tous leur importance (Berger et Luckmann, 1968).

Comme House et Howe nous le rappellent, les faits et les jugements de valeur se rejoignent sur un continuum, où l'évolution de l'institution d'évaluation ainsi que ses procédures, établies par des évaluateurs professionnels, déterminent si une affirmation est vraie et objective. Le problème réside donc dans les jugements de valeur qui ne sont pas rendus de manière explicite, tel qu'il est expliqué par House et Howe dans leurs études sur l'évaluation. Le problème de l'« incompréhension » entre les gestionnaires et les évaluateurs vient du fait que ces derniers prétendent à une impartialité, une objectivité et une neutralité de jugement. Ceci ouvre la porte aux critiques des gestionnaires en ce qui concerne le biais

des évaluations par rapport aux thèmes à la mode dans la communauté de développement. Comme nous le rapporte un interviewé, « des concepts tels que la participation, le développement durable, la bonne gouvernance, l'égalité des sexes et le mode de vie des populations autochtones ne sont que des termes en vogue dans la communauté de développement lesquels sont souvent utilisés exagérément par les évaluateurs dans leurs rapports ».

Il convient donc de se questionner sur les conditions qui donnent lieu à la formulation de jugements de valeur et à leur émergence dans le discours tenu par les évaluateurs. Trop souvent, les évaluateurs prennent pour acquis le fait qu'ils opèrent dans des sociétés et des organisations démocratiques où le débat, le dialogue ouvert et l'interaction avec les différents acteurs sont des principes partagés et acceptés par tous (House et Howe, 1999). Les évaluateurs devraient plutôt se questionner sur ce que les gens pensent en regard des politiques de développement et des initiatives de programmes. Ceci leur permettrait de voir et de comprendre comment les gestionnaires de programmes conçoivent le monde (House et Howe, 1999).

Parmi les quatre études de cas évaluées, deux seulement, celles sur l'industrie forestière et sur l'anti-corruption, questionnent de manière concrète les prémisses prescriptives à la base de la logique des programmes. Par conséquent, la fonction clarificatrice de l'évaluation n'a pas été entièrement remplie. Pour maximiser le processus, il s'avérerait plus opportun de s'interroger sur l'évolution des idées à la base des politiques et des programmes de développement de manière à comprendre quelle est l'opinion partagée sur ce qui mériterait d'être amélioré. Outre la culture, il convient de s'arrêter à la politique à la base du processus d'évaluation. Dans le prochain chapitre, nous verrons en quoi certaine information découlant du processus d'évaluation vient influencer sur les décisions et les actions subséquentes.

## ■ REMETTRE EN QUESTION OU LÉGITIMER L'ÉVALUATION ? LE DÉLAI NÉCESSAIRE DU CONSENSUS POLITIQUE

En accord avec une conviction scientifique largement partagée, des entrevues ont démontré que les changements de politique et de philosophie de la Banque suggérés par les évaluations sont trop souvent inutiles (Weiss, 1977, Weiss, Bucuvalas 1980a, 1980b; Caplan, 1975; Vedung, 1997). Pour les gestionnaires de programmes, l'information présente dans les rapports d'évaluation ne renferme pas de résultats pertinents et s'il y a de l'information pertinente, elle se doit d'être conforme aux attentes des utilisateurs pour qu'elle leur soit utile. Pour plusieurs répondants, l'étude de l'industrie forestière s'est avérée utile pour la conceptualisation de « l'effet de refroidissement ». Cette conceptualisation est cohérente avec la compréhension que les employés ont de la réalité des choses. De plus, elle s'avère représentative du désengagement de ces derniers vis-à-vis de ce type de projet étant donné le risque d'échec élevé. En bref, les employés souhaitent maintenir le *statu quo* dans leur prise de décision. Ils seront réceptifs à de nouvelles recommandations dans la mesure où elles s'inscrivent à travers des politiques et des philosophies qui leur paraissent acceptables (Weiss, Bucuvalas 1980a).

Selon l'influence du facteur de conformité établi par Weiss en 1981 dans son étude sur la santé mentale, les employés de la Banque filtrent l'information en fonction de leurs connaissances et de leurs croyances.

**Si quelqu'un souhaite vraiment prendre l'évaluation au sérieux, il doit être prêt à se battre, remettant ainsi en question sa propre carrière et celle de ses collègues.**

Trois des études de cas semblent démontrer que les employés de la Banque ont un cadre d'analyse et de compréhension déjà formé. En renforçant et en élaborant sur ce que les gestionnaires et les décideurs savent déjà de par leurs expériences, leurs lectures et leurs implications directes, certaines évaluations sont apparues intuitivement plausibles et ont ainsi pu être prises en considération. D'après les entrevues, les perceptions qui sont ressorties des cas de l'industrie forestière, de la construction de

barrages, et, dans une moindre mesure, de l'étude du PER et de l'anticorruption, ont semblé légitimer une position déjà partagée par plusieurs employés de la Banque.

**L'étude sur l'industrie forestière de l'OED est un abrégé, une synthèse riche d'une connaissance que nous avons déjà développée. Généralement, les études de l'OED présentent des critiques que les employés de la Banque ont déjà reçues.**

**Nous connaissons déjà cette information et il n'y a rien de nouveau dans le rapport sur l'industrie forestière. La valeur ajoutée de l'OED, c'est qu'il observe la situation d'un œil indépendant et présente une analyse systématique de la situation.**

**Les rapports de l'OED sont utiles dans la mesure où ils sont compatibles avec la position des personnes dans cette institution.**

**Les études de l'OED sont plausibles dans la mesure où leur contenu correspond à ma propre intuition sur le fonctionnement des choses. Si l'OED échoue à reconnaître le bon sens, les gens ne croiront pas que les barrages sont une bonne chose pour le développement.**

Tel qu'il est indiqué dans certaines entrevues, il importe en premier lieu de se questionner sur le tropisme que peuvent percevoir les employés face aux évaluations qui remettent en question le *statu quo*. En second lieu, il importe d'évaluer la possibilité d'une caractéristique méthodologique ou analytique qui limite la portée et la profondeur des évaluations.

**Les évaluations de l'OED sont technocratiques dans la mesure où elles cherchent à répondre à des questions précises (...) par exemple, en quoi le programme atteint ses objectifs. Les études de l'OED ne cherchent pas à évaluer des alternatives possibles à la situation présente (...) elles ne comparent jamais le plan A au plan B. Ceci n'implique pas le fait qu'il faille évaluer par exemple les villages avec fond de solidarité sociale et les villages sans fond de solidarité sociale (...) certainement ceci est déjà mieux que rien. (...) Mais la meilleure chose serait de formuler des questions qui tentent de trouver la meilleure solution possible.**

En effet, les évaluations observées dans nos études tentent de mesurer l'atteinte des objectifs en regard de certains critères (même si l'étude sur la lutte contre la corruption est une évaluation à mi-parcours),

avec une attention particulière sur la pertinence et l'efficacité des programmes d'assistance de la Banque ou la durabilité et l'impact des développements institutionnels. Bien qu'elles cherchent à évaluer différents aspects des programmes – conception, mise en œuvre, résultats et impacts – ces évaluations tendent toujours à mesurer la cohérence entre les résultats et les objectifs escomptés. L'objectif premier est d'évaluer la performance de la Banque.

L'analyse du changement des objectifs de développement en fonction de l'évolution des conditions sur le terrain – dans le but de concevoir les programmes indépendamment des institutions, des organisations, des relations socio-économiques et politiques – n'a été faite que dans les cas de l'étude de l'industrie forestière et de l'étude anticorruption. Dans les deux autres études de cas, l'attention accordée par les évaluateurs au contexte s'est avérée être une reconstruction historique plutôt qu'une analyse réelle du rôle de ce contexte. L'approche utilisée dans les évaluations ne laisse aucune place pour des considérations allant à l'encontre des faits ou permettant de percevoir des effets inattendus.

Ainsi, tout effort visant à caractériser l'évaluation comme remettant en question le *statu quo* ou comme s'inscrivant en ligne directe avec les attentes des acteurs donne lieu à une interprétation subjective. À cet égard, la théorie de la dissonance cognitive (Festinger, 1957) aide à expliquer le biais potentiel des gestionnaires. La théorie expose que l'effet principal d'une nouvelle information est de réduire la dissonance par rapport aux expériences, aux principes et aux connaissances propres à chacun. L'apprentissage n'est pas une conséquence de l'assimilation d'une nouvelle connaissance, laquelle vient influencer le comportement individuel et donc l'action organisationnelle.

L'apprentissage se fait plutôt dans le sens inverse: d'abord vient l'action ou le comportement, ensuite survient une modification de l'attitude et finalement la connaissance change afin de légitimer l'action posée. Il semble bien qu'il y ait un délai entre la présentation des résultats de l'évaluation et sa légitimation. La perception de menaces résultant de la remise en question du *statu quo* est graduellement rempla-

cée par une acceptation des résultats et des recommandations de l'évaluation. Quelque chose qui de prime abord apparaît contre nature devient progressivement acceptable. Spécialement en ce qui concerne le cas des PER, la résistance et l'opposition initiale ont progressivement disparu si bien que, deux ans plus tard, les conclusions et les recommandations de l'évaluation furent acceptées par tous. Néanmoins, il s'avère difficile d'attribuer cette acceptation au processus d'évaluation.

**Nous savons qu'après la lecture d'un rapport, le lecteur tente de minimiser ce qui y est écrit (...) Comment pouvons-nous savoir ce qui fut vraiment appris ?**

Cependant, selon certaines entrevues, des gestionnaires de programmes firent une distinction entre, d'une part, leurs propres valeurs et prises de position et, d'autre part, celles de l'institution à laquelle ils appartiennent. Selon ces témoignages, « l'acceptation politique » des évaluations représente un enjeu non pas pour les croyances et les préférences de l'individu, mais plutôt pour la vision et les processus opérationnels de l'organisation comme entité unitaire.

**Je serais très intéressée d'en savoir plus sur les avantages, pour un cours d'eau, de la construction de grands barrages, (...) d'avoir une meilleure compréhension de l'efficacité du développement de grands barrages.**

**S'ils me disent que les barrages sont une bonne chose, ils doivent en faire la preuve. Le fait est que dans cette institution, on ne construit pas de barrages, car les inspecteurs et les Organisations Non Gouvernementales nous observent de beaucoup trop près !**

Ces considérations suggèrent que les études qui confortent et rendent légitime le travail des gestionnaires de programmes ont de bonne chance d'être prises en compte, car elles reconnaissent la valeur ajoutée des évaluations en raison de leur appréciation systématique, de leur cadre d'analyse, de la collecte des données et de leur indépendance de jugement. Parmi les quatre études de cas, ces facteurs sont ressortis comme étant les caractéristiques de l'évaluation les plus utiles pour améliorer la stratégie organisationnelle et l'action.



La qualité de la méthodologie utilisée est aussi apparue comme un facteur augmentant indirectement la crédibilité de l'évaluation. Selon les répondants, bien que l'exactitude et la rigueur de la méthodologie aient été appréciées, les choix méthodologiques, de par leur nature trop technique, furent laissés aux évaluateurs. La méthodologie devient un enjeu dans le cas de questions particulières, comme lors de l'évaluation de grands barrages ou dans l'étude des PER. Dans ces cas, l'expertise s'avérait insuffisante pour permettre de résoudre des problèmes de développement des plus critiques.

Dans l'ensemble, nos études de cas ont démontré que la politique influence la procédure institutionnelle et le comportement institutionnel et organisationnel en relation avec l'utilisation faite de l'information des évaluations. Tant les gestionnaires de programmes que les décideurs ont admis l'utilité politique de l'évaluation dans l'optique où celle-ci permet de légitimer leur position ou leur vision des programmes. Dans ce cas, l'évaluation risque d'être incorporée dans la prise de décision future. Qui plus est, on semble accorder de l'importance à l'évaluation lorsque celle-ci touche un enjeu important et présent au programme politique du moment. Malheureusement, dès que l'attention est portée sur un autre enjeu, l'évaluation est aussitôt oubliée.

Dans cette perspective, la nature rétrospective propre au programme d'évaluation peut limiter sa contribution à la prise de décision en cours. Certains répondants ont mentionné l'importance de l'évaluation dans le cadre de l'étude sur l'industrie forestière. Celle-ci a permis de reformuler subséquemment la stratégie de cette industrie. Étant donné que le processus d'évaluation était inclus dans une politique élargie de révision de cette industrie, plusieurs conclusions présentes dans l'évaluation furent reçues par une audience réceptive qui a pu mettre immédiatement en pratique de telles recommandations. Mais dans le cas de l'étude sur la construction des barrages, même quatre ans après que le rapport ait été complété, certaines recommandations de l'évaluation concernant les standards de sécurité des barrages ont été utilisées pour avaliser certaines décisions.

## ■ AMÉLIORER LES RELATIONS ENTRE LA GESTION, L'ÉVALUATION ET LES STRUCTURES DE GOUVERNANCE

La majorité des recherches portant sur le comportement organisationnel et l'administration publique ont étudié des structures de gouvernance qui contraignent et permettent l'action des gestionnaires publiques (Feldman et Kademiam 2002; Kettl, 2000, 2002; Behn, 2001; Acouin, 1990). Selon ces études, les décideurs imposent des structures de gouvernance plutôt statiques et ils rendent responsables les gestionnaires des résultats escomptés. En d'autres termes, les gestionnaires publics sont affectés par les structures qui les gouvernent et sont désavantagés par rapport à l'influence qu'ils peuvent exercer sur ces mêmes structures. Les processus d'évaluation et les mesures de performance sont des structures de gouvernance qui fournissent un mécanisme automatique, lequel permet l'exercice du contrôle et de la reddition de comptes.

Nos découvertes semblent démontrer que les structures de gouvernance ne sont pas toujours responsables de la performance organisationnelle. Les relations de travail qui prévalent entre les évaluateurs et les gestionnaires présentent des caractéristiques qui sont déterminées plus par la forme institutionnelle de l'organisation, ses coutumes et ses pratiques que par des facteurs externes à la Banque. Le cadre institutionnel de la fonction d'évaluation de la Banque mondiale encourage l'utilisation des évaluations pour la prise de décisions. Plus précisément, la position centrale et l'indépendance de l'évaluation, sa responsabilité devant le Conseil et son budget alloué par le Conseil et non par les gestionnaires sont des caractéristiques qui semblent prévenir les biais lors de l'évaluation des programmes. L'organisation fonctionnelle de l'évaluation permet également de développer les capacités techniques et aussi de renforcer l'acceptation et l'intégration de l'évaluation à travers l'ensemble de l'organisation.

En pratique, les liens organisationnels entre l'évaluation et les autres fonctions ne sont pas univoques (Goodman, 2000). Le rôle politique important inhérent au processus d'évaluation aide cette dernière à être acceptée au cœur même de l'organisation.

Pour que l'évaluation joue un rôle clef dans la prise de décision stratégique, le leadership démontré par le département d'évaluation est essentiel. Dans deux des quatre études de cas, l'évaluation influence de manière non négligeable la planification stratégique, informe sur les expériences passées de la Banque mondiale et permet d'opérer un équilibre entre les besoins de gestion et les priorités du conseil.

En revanche, il apparaît que le lien organisationnel entre l'évaluation et la gestion est plus faible au niveau inférieur de l'organisation (Goodman, 2000). Dans une certaine mesure, l'évaluation est perçue comme une activité des employés permettant la gestion axée sur les résultats. Le contenu analytique et technique de l'évaluation ainsi que son processus méthodologique – construit sur des analyses statistiques, des processus de consultation et sur des recommandations basées sur des faits – font que les rapports d'évaluation s'apparentent davantage à un travail de conseil et d'analyse. Ce qui fait que l'évaluation est différente, c'est son regard systématique sur les expériences passées et sur la fonction organisationnelle de « contrôle et équilibre » à l'intérieur de l'organisation.

Les barrières culturelles et politiques empêchent la mise en place d'une interaction effective entre les gestionnaires et les évaluateurs pour ce qui est de l'échange d'information et de la confiance mutuelle, particulièrement au niveau le plus bas de l'organisation. Le manque d'enthousiasme pour l'information dérivant des évaluations fait montre d'une culture hiérarchique, basée sur la commande et le contrôle, laquelle ne permet pas d'exploiter le potentiel de l'évaluation comme moteur de changement organisationnel et comme réflexion critique sur tous les plans de la gestion.

Qui plus est, l'évaluation est soumise aux valeurs et à la culture organisationnelle – qui influencent la sélection des évaluations – et au programme du département d'évaluation. La composante normative de l'évaluation qui agit comme source de communication entre les gestionnaires et les évaluateurs fonctionne dans la mesure où les points de vue, les valeurs, les idées et les connaissances à la base des politiques de développement n'apparais-

sent pas de manière flagrante. Enfin, le délai d'acceptation concernant les recommandations de l'évaluation est à mettre en relation avec le consensus politique. Celui-ci permet à une idée auparavant inacceptable de devenir progressivement acceptable aux yeux des gestionnaires.

Ces résultats signifient que le lien organisationnel entre l'évaluateur et le gestionnaire, dans leurs relations quotidiennes, dépend de l'échange lors de la conception de l'évaluation, de la collecte des données, de l'analyse et de la transmission de données. Pour que cet effort porte ses fruits, il importe d'intégrer les différents membres de l'organisation et les parties intéressées, et ce, à toutes les étapes du processus d'évaluation. Avec le temps, les interactions systématiques prenant place à travers des canaux informels permettront aux évaluateurs et aux gestionnaires de comparer leurs expériences et leurs pratiques de travail, et, ultimement, d'adopter la connaissance pratique dérivant du processus d'évaluation.

En somme, l'organisation institutionnelle et les structures de gouvernance, mais aussi la capacité de développer des relations formelles et informelles qui promeuvent l'échange d'information et de connaissances restent essentielles au bon fonctionnement de l'évaluation. En d'autres mots, l'influence de l'évaluation sur les décisions organisationnelles dépend aussi de la faculté qu'ont les évaluateurs de mettre en place des études dont les procédures permettront la participation des gestionnaires de programmes. Bien que de meilleures structures, règles et procédures organisationnelles améliorent le fonctionnement de l'évaluation à l'intérieur de l'organisation, il est essentiel que les évaluateurs et les gestionnaires contribuent ensemble à créer, changer et maintenir ces structures dans le but de maximiser le fonctionnement organisationnel de la Banque mondiale.

## BIBLIOGRAPHIE

- Argyris, C. (1994). *On Organizational Learning*, Malden, Blackwell.
- Aucoin, P. (1990). « Administrative Reform in Public Management », *Governance*, vol. 3, n° 2.
- Barzelay, M. (1992). *Breaking through Bureaucracy: A New Vision of Management in Government*, University of California Press.
- Barzelay, M. (2001). *The New Public Management: Improving Research and Policy Dialogue*, University of California Press.
- Behn, R. D. (2001). *Rethinking Democratic Accountability*, The Brookings Institution.
- Berger, P., T. Luckmann (1968). *The Social Construction of Reality*, New York, Doubleday.
- Boyle, R., D. Lemaire, Eds. (1999). *Building Effective Evaluation Capacity. Lessons from Practice*. Transaction Publishers, New Brunswick (USA) and London (UK).
- Bresser Pereira, L. C. (1999). *Reforming the State. Managerial Public Administration in Latin America*, Londres, Lynne Rienner Publishers.
- Caplan, N. (1975). "The Use of Social Science Information by Federal Executives." dans Lyons (ed.), *Social Science and Public Policies*, The Dartmouth/OECD Conference, Hanover, pp. 47-67.
- Chelimsky, E., W. Shadish (1997). *Evaluation for the 21st Century*, Thousands Oaks, Londres, New Delhi, Sage Publications.
- Feldman, M., A. Kademian (2002). « To Manage Is to Govern », *Public Administration Review*, vol. 62 n° 5.
- Festinger, L. (1957). *The Theory of Cognitive Dissonance*, New York, Harper-Collins.
- Gray, A., B. Jenkins, B. Segsworth, R. C. Rist (1993). *Budgeting, Auditing, and Evaluation: Functions and Integration in Seven Countries*, Transaction Publishers.
- Gore, A. (1993). *Putting Customers First. How to Create a Government that Works Better and Costs Less*, National Performance Review.
- Goodman, P. S. (2000). *Missing Organizational Linkages: Tools for Cross-Level Research*, SAGE Publications.
- House, E. R., K.R. Howe (1999). *Values in Evaluation and Social Research*, Thousand Oaks, Londres, SAGE Publications.
- International Financial Institutions Advisory Commissions (1998), Meltzer Commission Final Report, US Congress, Washington DC.
- Kapur, D., P.J. Lewis, R. Webb (Eds.) (1997). *The World Bank: Its First Half Century*, Washington DC, The Brookings Institution.
- Kettl, D. (2000). *The Global Public Management Revolution: A Report on the Transformation of Governance*, The Brookings Institution.
- Kettl, D. (2002). *The Transformation of Governance: Public Administration for Twenty-First Century*. America, Johns Hopkins University Press.
- Kirkhart, K.E. (2000). « Reconceptualizing Evaluation Use: An Integrated Theory of Influence », *New Directions for Evaluation*, n° 88, hiver, Jossey-Bass Publishers.
- Knott, J. H., G. J. Miller (1998). *Reforming Bureaucracy: The Politics of Institutional Choice*, Prentice Hall.
- Marra, M. (2000). « How Much Does Evaluation Matter? Some Examples of Utilization of World Bank's Anticorruption Activities' Evaluation », *Evaluation*, vol. 6, n° 1.
- Meyer, J. W., B. Rowan (1991). « Institutionalized Organizations: Formal Structure as Myth and Ceremony », dans Powell, W. W. et P. J. DiMaggio (Eds.) (1991). *The New Institutionalism in Organizational Analysis*, Chicago et Londres, The University of Chicago Press.
- Miller, G.J. (1992). *Managerial Dilemmas. The Political Economy of Hierarchy*, Cambridge University Press.

- Moore, M. H. (1998). *Creating Public Value. Strategic Management in Government*, Cambridge, Massachusetts, Londres, Angleterre, Harvard University Press.
- Newcomer, K. E. (2001). «Tracking and Probing Program Performance: Fruitful Path or Blind Alley for Evaluation Professionals», *The American Journal of Evaluation*, vol. 22 n° 3, p. 343-348.
- Nioche J. P., R. Poinard (1984). *L'Évaluation des politiques publiques*, Paris, Economica.
- Nonaka, I. (1994). «A Dynamic Theory of Organizational Knowledge Creation», *Organization Science*, vol. 5, n° 1.
- North, D. (1990). *Institutions, institutional change, and economic performance*, Cambridge University Press.
- OCDE, (1995). *Governance in Transition*, Paris, OCDE.
- OCDE (1996). *Principles for Evaluation of Development Assistance*, Development Assistance Committee (DAC).
- Operations Evaluation Department (1996). *The World Bank's Experience with Large Dams – A Preliminary Review of Impact*, Washington DC, World Bank.
- Operations Evaluation Department (1998a). *Public Sector Performance – The Critical Role of Evaluation*, Selected Proceedings from a World Bank Seminar, Washington DC, World Bank.
- Operations Evaluation Department (1998b). *Recent Experience With Involuntary Resettlement. Thailand – Pak Mun*, Report N° 17541, Washington DC, World Bank.
- Operations Evaluation Department (1998c). *The Impact of Public Expenditure Review*, Washington DC, World Bank.
- Operations Evaluation Department (1998d). *Assessing Development Effectiveness*, Washington DC, World Bank.
- Operations Evaluation Department (2000a). «A Review of the World Bank's 1991 Forest Strategy and its Implementation», Volume I, Main Report, Washington DC, World Bank.
- Operations Evaluation Department (2000b). *The World Bank Forestry Strategy – Striking the Right Balance*, Washington DC, World Bank.
- Operations Evaluation Department (2000c). *Glossary of Evaluation Terms*, internal methodology paper, Washington DC, World Bank.
- Osborne, D., T. Gaebler (1993). *Reinventing Government. How the Entrepreneurial Spirit Is Transforming the Public Sector*, Plume.
- Picciotto, R. (1997). «Evaluation in the World Bank: Antecedents, Instruments, and Concepts» dans Chelimsky, E. et W. Shadish (Eds.) (1997). *Evaluation for the 21<sup>st</sup> Century*, Thousands Oaks, Londres, New Delhi, SAGE Publications.
- Picciotto, R., E. Wiesner (1998). *Evaluation and Development, The Institutional Dimension*. Transaction Publishers, New Brunswick (USA) and London (UK).
- Picciotto, R. (2002). "Evaluation at the World Bank: Antecedents, Methods, and Instruments" dans Rist, R.C., J.E. Furubo, (2002), *International Atlas of Evaluation*, Transactions Publishers, New Brunswick (USA) and London (UK).
- Pollitt, C. (1995). «Justification by Works or by Faith?», *Evaluation*, vol. 1, n° 2, p. 133-154.
- Powell, W. W., P. J. DiMaggio (Eds.) (1991). *The New Institutionalism in Organizational Analysis*, Chicago et Londres, The University of Chicago Press.
- Scott. R. W. (1991). «Unpacking Institutional Arguments», dans Powell, W. W. et P. J. DiMaggio (Eds.) (1991). *The New Institutionalism in Organizational Analysis*, Chicago et Londres, The University of Chicago Press.
- Svensson, K. (1997). «The Analysis and Evaluation of Foreign Aid», dans Chelimsky, E. et W. Shadish (Eds.) (1997). *Evaluation for the 21<sup>st</sup> Century*, Thousands Oaks, Londres, New Delhi, SAGE Publications.
- Thoenig, J.C. (1999). *Evaluation as Usable Knowledge for Public Management Reforms*, PUMA/SGF, Symposium on «Government of the Future», Paris.
- Vedung, E. (1997). *Public Policy and Program Evaluation*, Transaction Publishers.
- Yin, R. K. (1994). *Case Study Research: Design and Methods*, Second Edition, Applied Social Research Methods Series, vol. 5, Thousand Oaks, London, New Delhi, SAGE Publications.
- Weiss, C. H. (Ed.), (1977). *Using social research in public policy making*. Lexington, Mass., Lexington Books.

- Weiss, C., M. Bucuvalas (1980a). *Social Science Research and Decision Making*, Columbia University Press.
- Weiss, C., M. Bucuvalas (1980b). « Truth Tests and Utility Tests: Decision Makers' Frames of Reference for Social Science Research » *American Sociological Review*, vol. 45, n° 1, p. 302-313.
- Weiss, C.H. (1997). *Evaluation. Methods for Studying Programs and Policies*. Second Edition, Prentice Hall, Upper Saddle River, New Jersey.
- Wholey, J. S. (2001). « Managing for Results: Roles for Evaluators in a New Management Era » *The American Journal of Evaluation*, vol. 22, n° 3, p. 343-348.
- Williamson, O. (1975). *Markets and Hierarchies*, New York, Free Press.
- World Commission on Dams (2000). *Dams and Development: A New Framework for Decision-making*, Londres et Sterling VA, EarthScan Publications Ltd.
- Zucker, L. G. (1991). « The Role of Institutionalization in Cultural Persistence » dans Powell, W. W. et P. J. DiMaggio (Eds.) (1991). *The New Institutionalism in Organizational Analysis*, Chicago et Londres, The University of Chicago Press.

# ENAP

## L'UNIVERSITÉ DE L'ADMINISTRATION PUBLIQUE

www.enap.ca

L'École nationale d'administration publique forme, depuis 35 ans, les cadres et les professionnels du secteur public, grâce à un **enseignement crédité de 2<sup>e</sup> et de 3<sup>e</sup> cycles** et à **des activités de perfectionnement**.

L'ENAP contribue en outre au développement de l'administration publique par des activités de **recherche** et par des services **d'appui aux organisations**.

Comme partenaire universitaire de l'administration publique, elle oriente ses activités de **recherche** dans les domaines suivants :

- l'évaluation et le développement des compétences
- l'évaluation des politiques et des programmes publics
- l'éthique publique
- le leadership
- l'impact de la mondialisation sur les politiques publiques
- les technologies et l'organisation du travail
- la gouvernance

L'**Observatoire de l'administration publique** de L'ENAP offre aux décideurs publics un comparatif stratégique en matière de gestion et de politiques publiques. L'Observatoire diffuse et publie les résultats de ses recherches dans plusieurs publications, dont *Télescope*.

Sur le plan international, l'ENAP dispense des programmes d'études en ligne et des programmes de formation continue. Ses interventions prennent la forme d'appui institutionnel aux établissements de formation en administration publique et de services conseils auprès des gouvernements en matière de gouvernance et de réforme du secteur public.

L'ENAP est reconnue au Québec et dans le monde comme l'une des plus importantes écoles universitaires de management public de langue française. Elle accueille annuellement autour de 1 700 étudiants et compte plus de 5 500 diplômés, dont plusieurs exercent des fonctions stratégiques dans le secteur public.



Université du Québec

**École nationale d'administration publique**

# L'évaluation des programmes d'aide au développement conduits par la **Commission européenne**. Les grandes synthèses thématiques et par pays

Par **Luc Lefebvre**, administrateur délégué de la société d'études et de conseil ADE<sup>1</sup>,  
enseignant à l'Institut d'études politiques (IEP – Sciences Po) de Paris • luc.lefebvre@ade.be  
Et **Ivory Yong**, responsable du département d'évaluation de la société ADE • ivory.yong@ade.be

L'aide européenne aux pays en développement est importante et diversifiée. Le cadre institutionnel dans lequel elle s'exerce, le contexte historique de ces 15 dernières années et l'évolution des modes de gestion de la Commission européenne ont contribué à favoriser le développement rapide de l'évaluation de cette aide et l'expression plus formelle des priorités politiques qui la sous-tendent. Ces politiques, exprimées au niveau sectoriel ou à celui des stratégies de coopération avec chaque pays, font l'objet d'évaluations de synthèse qui posent des problèmes méthodologiques difficiles. Une évolution de l'approche évaluative vers des méthodes plus comparatives et participatives mériterait d'être explorée.

## ■ UNE AIDE AU DÉVELOPPEMENT IMPORTANTE ET DIVERSIFIÉE

L'Union Européenne (UE) a consacré 34,3 milliards d'euros à l'aide au développement en 2004. Cette aide représente plus de la moitié de l'aide publique au développement enregistrée par le CAD<sup>2</sup>. Environ 20 % de ce montant, soit près de 7 milliards d'euros, constitue l'aide communautaire, gérée par la Commission européenne. Le reste est géré directement par les États membres.

L'aide européenne aux pays tiers<sup>3</sup> vise à lutter contre la pauvreté, à favoriser l'intégration des économies des pays partenaires dans l'économie mondiale et à promouvoir un développement durable. Elle s'inscrit dans une politique extérieure de l'UE qui soutient des valeurs telles que le respect des droits de l'homme, la promotion de l'État de droit, la démocratie et la bonne gouvernance. Les objectifs sont donc nombreux et appliqués avec des accents variables selon les régions, comme en témoignent les accords de coopération spécifiques signés avec chacune d'entre elles.

<sup>1</sup> www.ade.be

Les auteurs ont contribué à certaines des évaluations décrites dans le présent texte et ont eu une influence sur les choix méthodologiques qui sous-tendent la démarche. Leur regard n'est donc pas celui d'observateurs indépendants, mais plutôt celui de praticiens directement impliqués dans l'action.

<sup>2</sup> Comité d'aide au développement de l'OCDE.

<sup>3</sup> Pour simplifier, on entendra ici par « aide européenne aux pays tiers » l'aide gérée par la Commission européenne aux pays n'appartenant pas à l'UE, à l'exclusion des pays ayant entamé un processus d'accession à l'Union.

Dans les pays d'Afrique, des Caraïbes et du Pacifique (ACP), liés à l'Europe par une convention de coopération pluriannuelle, l'aide européenne a représenté 2,6 milliards d'euros en 2004<sup>4</sup>. Dans cette région, la réduction de la pauvreté est clairement affichée comme étant la priorité. L'intégration de ces pays dans l'économie mondiale est perçue comme un moyen de contribuer à ce premier objectif. Dans les pays d'ACP, l'Europe, un des principaux bailleurs de fonds, est impliquée dans la plupart des secteurs et dans les principaux choix de politique économique des pays bénéficiaires.

La coopération avec les pays méditerranéens et du Moyen-Orient, à laquelle l'aide européenne a consacré 700 millions d'euros en 2004, s'inscrit dans une politique dite de voisinage. L'ambition est notamment de créer une zone de libre-échange entre l'Europe et les pays du sud de la Méditerranée. Les enjeux commerciaux, mais aussi la question des flux migratoires et l'interconnexion des réseaux de communication constituent des dimensions importantes de cette coopération.

Dans l'ouest des Balkans (ex-Yougoslavie), l'Europe a engagé 663 millions d'euros en 2004, pour soutenir le processus de pacification et de reconstruction d'un espace politique stable et économiquement viable. La référence aux « acquis communautaires » et donc la perspective d'une accession à l'UE sous-tendent la démarche.

En Europe de l'Est et en Asie centrale, un programme européen d'assistance technique de 504 millions d'euros en 2004, poursuit des objectifs tels que l'appui à la transition vers l'économie de marché, la coopération transfrontalière et la sécurité nucléaire.

En Asie et en Amérique latine, l'aide européenne, malgré un volume total de 923 millions d'euros en 2004, joue un rôle moins important, tant par rapport aux ressources propres de ces pays que par rapport aux autres bailleurs de fonds. Dans ces régions, à quelques exceptions près telles que la Bolivie, le Bangladesh ou le Vietnam pour lesquels

l'aide européenne a un poids significatif, l'Europe vise essentiellement à assurer une présence et à appuyer quelques opérations particulières. Les conventions de coopération avec les pays asiatiques et latino-américains se réfèrent à la notion d'intérêt mutuel qui traduit l'esprit de la coopération européenne dans ces régions.

Malgré cette diversité, la Commission européenne a formulé sa politique de développement sur l'ensemble des pays tiers. Un bref survol historique rappelle que cette nécessité de cohérence politique et le développement de la pratique de l'évaluation sont allés de pair.

### ■ ÉVALUER L'AIDE, UNE PRATIQUE RÉCENTE MAIS INCONTOURNABLE

Au cours de la dernière décennie, trois facteurs convergents ont été à la base d'un développement important de l'évaluation de l'aide européenne : a) le cadre institutionnel particulier dans lequel s'exerce l'aide internationale; b) l'opportunité politique ouverte à la fin de la guerre froide; c) la tendance générale vers une gestion de la Commission européenne plus orientée vers les résultats.

#### L'aide internationale, un cadre institutionnel particulier

Le contrôle de l'action publique par le citoyen est un enjeu crucial de nos démocraties. Ce contrôle, on le sait, est d'autant plus difficile à exercer que, d'une part, la nature de l'action publique se complexifie et, d'autre part, nos sociétés sont plus hétérogènes dans le sens où elles se composent de groupes défendant des systèmes de valeurs de plus en plus contrastés. En outre, certaines valeurs défendues de longue date prennent une dimension différente. On pense notamment au désarroi des promoteurs du discours « rhénan » sur la solidarité sociale et la régulation du marché du travail face, non seulement à la concurrence mondiale, mais également à l'évolution de la notion même de soli-

<sup>4</sup> Les chiffres cités proviennent du Rapport annuel sur la politique de développement de la Communauté européenne, Commission européenne (2005).



darité. Le facteur 100 qui distingue ce que nous donnons aux « nôtres » de ce que nous donnons aux « autres » pose en effet une question éthique de plus en plus difficile à occulter. Les Européens, pourtant réputés généreux, on vient de le voir, consacrent plus de 30% de leurs revenus à leurs systèmes nationaux de solidarité sociale et autour de 0,3% à la solidarité internationale. Nous donnons peu à ceux qui ont commis l'erreur de naître du mauvais côté de nos frontières et une des raisons souvent avancées de notre parcimonie est que nous ne savons pas très bien « ce qu'ils font de cet argent ».

Le contrôle de l'action publique par le citoyen s'exerce dans des conditions très différentes selon que cette action se réalise au sein de l'espace national ou dans un pays tiers. En effet, comme l'indique la figure 1, au sein de l'espace national, le citoyen exerce en tant qu'électeur un contrôle par l'entremise du parlement sur les décisions politiques prises par l'exécutif. En tant que contribuable, il fournit à ce dernier les moyens de son action et il est lui-même bénéficiaire (parfois victime) des actions entreprises par le gouvernement, que celui-ci les conduise directement par son administration, ou indirectement en s'appuyant sur les entreprises ou d'autres formes d'association. En principe, le ci-

toyen est à l'origine des choix politiques, en finance la mise en œuvre et en voit les résultats. On dira dans ce cas que la boucle de rétroaction est fermée. Une telle configuration n'implique évidemment pas que l'information du citoyen, ni même celle des acteurs de l'intervention publique, soit parfaite. L'évaluation a certes un rôle à jouer dans ce cadre comme moyen d'éclairage critique de l'action à la fois dans la perspective de permettre aux décideurs de tirer les leçons de l'expérience (perspective formative), mais également pour rendre compte au citoyen qui, aux divers titres évoqués ci-dessus, peut, à travers ses représentants, exercer un pouvoir (perspective sommative).

La configuration institutionnelle de l'aide internationale est toutefois très différente (figure 2). Dans ce cas, le citoyen-contribuable et le citoyen-bénéficiaire appartiennent à deux espaces politiques distincts. Même dans le cas le plus favorable où ces deux espaces sont démocratiques, le contribuable n'a que très peu de moyens d'apprécier l'effet des interventions qu'il finance et le bénéficiaire n'a le pouvoir politique d'influencer ni la fraction des impôts consacrée à l'aide, ni même souvent l'affectation de celle-ci.

En outre si, comme on l'a évoqué ci-dessus, le contrôle de l'action publique est plus difficile dans

FIGURE 1 : SCHÉMA INSTITUTIONNEL DE L'ACTION PUBLIQUE NATIONALE

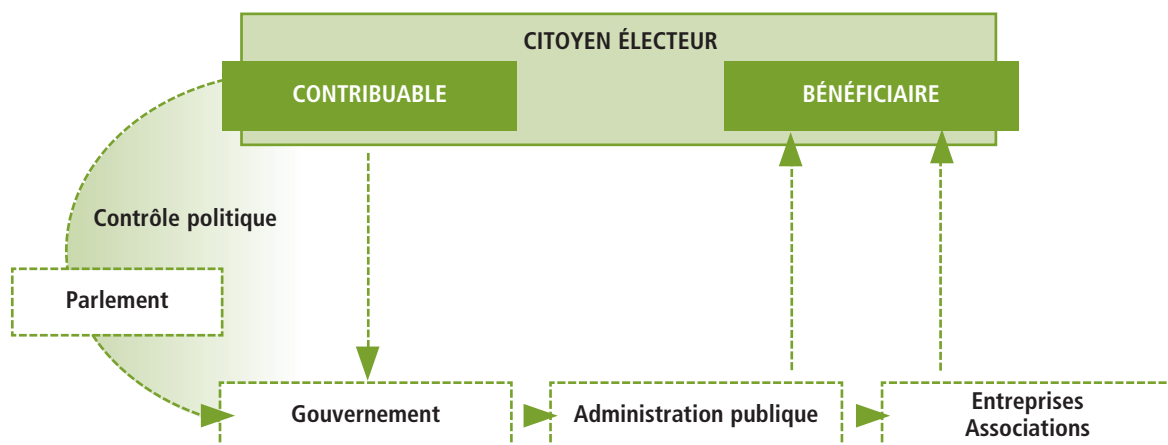
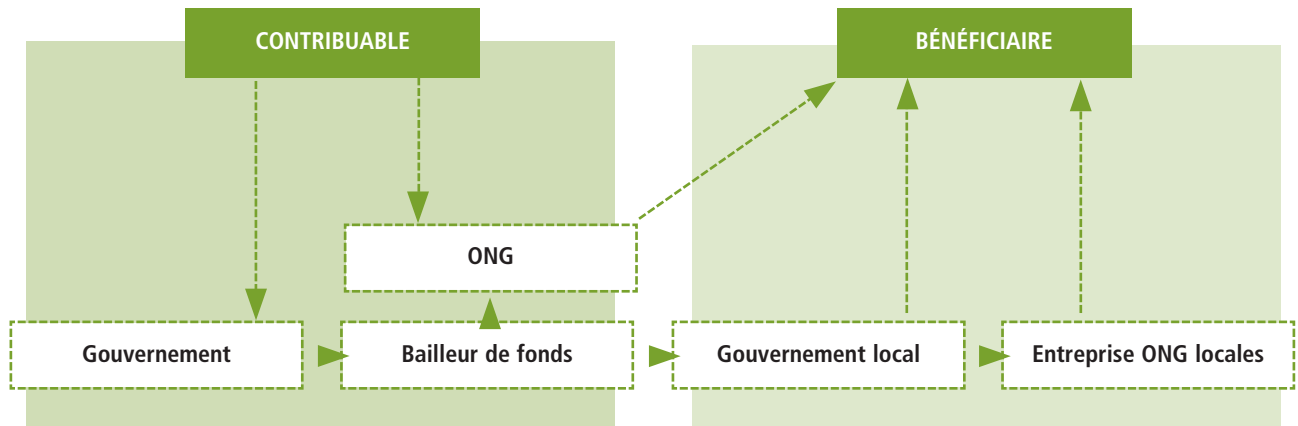


FIGURE 2 : SCHÉMA INSTITUTIONNEL DES INTERVENTIONS D'AIDE PUBLIQUE INTERNATIONALE



des sociétés composées de groupes très différenciés en termes de systèmes de valeurs, dans le cas de l'aide internationale, la situation est plus complexe encore. En effet, les différences de statuts économiques, d'univers culturels, de contextes historiques, de niveaux de formation séparant le contribuable et le bénéficiaire sont généralement telles que leur simple mise en contact ne suffirait probablement pas à établir entre eux la complicité nécessaire pour pouvoir contrôler les nombreux intermédiaires chargés de transférer les ressources de l'un vers l'autre. En d'autres termes, dans le contexte actuel, les intérêts et les valeurs des uns sont généralement trop différents de ceux des autres pour qu'une telle perspective paraisse réaliste. On parle de *boucle de rétroaction ouverte*, laissant les acteurs de la chaîne de distribution de l'aide quasiment sans contrôle démocratique<sup>5</sup>.

L'évaluation peut-elle contribuer à fermer cette boucle? En principe, l'évaluation peut effectivement jouer plusieurs rôles dans le schéma décrit ci-dessus: informer le bénéficiaire, informer le contribuable et permettre aux acteurs de tirer les leçons de leur expérience.

Informé le bénéficiaire contribue à sa capacité éventuelle d'exercer un contrôle sur le rôle des acteurs locaux. Relayer sa perception des effets de l'aide peut lui fournir indirectement les moyens d'en influencer l'affectation. La tendance est à l'appropriation de l'évaluation par les gouvernements et les autres acteurs des pays bénéficiaires, même si ce processus est limité par un manque de capacités. Si cette tendance se poursuit aux dépens du droit de regard du « contribuable-électeur », les partenaires occidentaux de l'aide risquent de fonctionner sans aucun contrôle, car les bénéficiaires et même leurs gouvernements ne disposent que de très peu de pouvoir sur les bailleurs de fonds et autres organisations non gouvernementales (ONG) internationales. Cette absence de contrôle démocratique de la chaîne d'acteurs de l'aide peut induire une perte de crédibilité de celle-ci et une diminution des financements. Les agences des Nations Unies constituent un exemple intéressant d'institutions dans lesquelles le pouvoir des gouvernements bénéficiaires est relativement important, mais qui éprouvent les plus grandes difficultés à financer leurs activités sans l'aide d'autres bailleurs tels que

<sup>5</sup> Lire à ce sujet Martens et al. (2002).

la Banque mondiale ou les agences bilatérales dont la structure de pouvoir est très différente.

Informar le contribuable des effets de son aide est indispensable afin de maintenir une pression suffisante sur les acteurs (bailleurs de fonds, ONG, entreprises) pour que ceux-ci se soucient de l'efficacité de leur intervention. Dans le contexte européen, le contribuable est représenté par les ministres des finances des États membres. Lorsque l'on sait que le principal canal d'aide européenne à l'Afrique est le Fond européen de développement, alimenté sur une base volontaire par les États membres, on mesure mieux l'importance de ce rôle.

Dans une perspective plus « formative », on observe des boucles plus courtes liant entre eux les acteurs de l'aide (bailleurs de fonds, gouvernements, ONG, entreprises). L'évaluation peut les aider à tirer les leçons du passé pour améliorer la qualité de leur action future. La nature de cette action et les objectifs qu'elle poursuit ont évolué dans le temps.

### **Une fenêtre politique favorable à l'analyse des effets de l'aide**

Du début des années 1960 jusqu'au début des années 1990, l'heure était à la décolonisation et à la guerre froide. Au début de cette période, l'aide au développement servait à protéger les intérêts des anciennes puissances coloniales, puis, progressivement, à maintenir les régimes assistés solidaires (captifs?) de ce que l'on appelait à l'époque le Monde libre. Dans ce contexte, la pertinence de l'aide paraissait stratégiquement évidente et son efficacité se mesurait en termes géopolitiques.

L'effondrement du bloc soviétique suscita de profondes remises en question. L'enjeu géopolitique que représentaient les pays « pauvres » parut nettement moins présent et les flux d'aide auraient pu tarir. Le risque était d'autant plus grand que lorsque l'on a réellement commencé à se demander

quelle était l'efficacité de l'aide en termes, notamment, d'appui à la croissance économique des pays bénéficiaires, les conclusions des premières études se révélèrent dévastatrices<sup>6</sup>. Il était urgent d'en savoir plus sur l'utilité de l'aide au développement. Une fenêtre politique s'était ouverte durant laquelle il était devenu non seulement possible d'évaluer l'aide par rapport au discours qui la justifiait, mais cette démarche était devenue nécessaire pour entretenir la volonté des citoyens ou de leurs représentants de contribuer au financement de ces interventions. L'évaluation, à la méthodologie encore balbutiante à l'époque, surtout s'agissant d'opérer dans des environnements très pauvres en données, connut un essor sans précédent. On évalua non seulement les projets individuels, mais des ensembles d'interventions par secteur ou par pays. Les premiers résultats ne contribuèrent pas vraiment à rassurer, sauf sur un point important : on commençait à se soucier de la pertinence, de l'efficacité, voire même de l'efficience des interventions de l'aide internationale.

Dès les premières évaluations sectorielles, conduites par la Commission européenne et d'autres bailleurs de fonds, la nécessité d'inscrire les interventions dans un cadre politique plus cohérent, basé sur une analyse plus systématique, apparut comme une recommandation majeure<sup>7</sup>. Parallèlement, et largement sous l'influence de ces évaluations, le discours stratégique des bailleurs de fonds s'est considérablement affiné et la coordination entre eux s'est améliorée. Des valeurs telles que la démocratie, les droits de l'homme, l'environnement et l'égalité des chances furent plus clairement affirmées comme des valeurs que l'aide internationale visait à promouvoir. La réduction de la pauvreté est devenue l'objectif général explicite de l'aide, fournissant au passage, pour un temps, l'aune commune et ultime dont les évaluateurs avaient besoin pour en apprécier les performances. La Commission européenne exprima très clairement ce choix en 2000 dans une Communication<sup>8</sup> puis dans une

<sup>6</sup> Lire à ce sujet diverses études lancées dès 1995 et synthétisées par Dollar et Pritchett (1998).

<sup>7</sup> Voir par exemple les évaluations réalisées dans les secteurs du développement rural (ADE, 1995) et des transports (Tecnecon, 1995).

<sup>8</sup> Commission européenne (2000a).

Résolution du Conseil<sup>9</sup>. Cette dernière constitua le premier document de politique d'aide au développement applicable à l'ensemble des pays bénéficiant de l'appui de l'Europe.

À la fin de 2001, la « guerre contre le terrorisme » referma brutalement, mais pas complètement, cette fenêtre. Les considérations géopolitiques reprirent le dessus dans bien des pays et leur poids dans les décisions d'allocations de ressources de l'aide redevint déterminant. Déterminant, mais probablement moins écrasant que lors de la guerre froide. À la Commission européenne, comme chez la plupart des bailleurs de fonds, notamment sous la pression d'une opinion publique plus sensibilisée<sup>10</sup>, l'évaluation ne perdit pas trop de ses acquis des années 1990. Elle continua au contraire à évoluer, partiellement supportée par une dynamique plus large, sans rapport direct avec l'aide au développement : la volonté affichée de la Commission européenne dans son ensemble, de faire évoluer ses pratiques d'une culture de gestion des ressources vers une gestion axée sur les résultats dans tous les domaines qu'elle couvre.

### **Une Commission européenne plus axée sur les résultats**

Depuis le début des années 2000, la Commission européenne a considérablement renforcé le statut de l'évaluation, considérant cette démarche comme une des clefs de voûte de la réorientation de ses méthodes de gestion vers ce que l'on appelle à Bruxelles la gestion par activité. L'objectif est de mettre l'accent sur les résultats plutôt que sur les ressources et les structures. Cela implique de met-

tre en place un système d'information et d'analyse adapté. Selon les communications présidant à cette réforme<sup>11</sup>:

- les activités de la Commission devront faire l'objet d'évaluations régulières, au moins une fois tous les six ans;
- les Directions générales (DG) devront avoir accès à des capacités d'évaluation spécialisées, internes ou en partage avec d'autres DG;
- une DG sera chargée de la coordination et du développement des capacités de la Commission en cette matière;
- on veillera à assurer l'intégration des conclusions des évaluations dans les décisions de programmation.

On se trouve donc devant un programme ambitieux et complet visant non seulement à développer des capacités en évaluation, mais à en généraliser la pratique et à s'assurer que les évaluations nourrissent réellement la décision. Pour les raisons évoquées ci-dessus, les DG en charge de l'aide au développement avaient de l'avance sur les autres et se sont donc inscrites assez facilement dans cette nouvelle approche.

---

<sup>9</sup> Commission européenne (2000d).

<sup>10</sup> « Une enquête Eurobaromètre spéciale, réalisée en 2004, confirme que la population européenne attend de l'Union qu'elle joue un rôle central en matière de développement. Six citoyens européens sur dix estiment que la Commission européenne aide les populations démunies des pays en développement, un objectif jugé important par neuf personnes sur dix. » (Commission européenne, 2005).

<sup>11</sup> Les réformes évoquées ont donné lieu à la production de nombreuses déclarations officielles, parmi lesquelles l'une des plus explicites à propos de l'évaluation est celle intitulée « Mettre l'accent sur les résultats : renforcer l'évaluation des activités de la Commission » (Commission européenne, 2000c).

## ■ LES GRANDES ÉVALUATIONS DE SYNTHÈSE

### Une responsabilité clairement définie

Depuis janvier 2001, trois DG de la Commission européenne sont directement impliquées dans la gestion de l'aide extérieure: la DG Développement, la DG Relations Extérieures et EuropeAid. Les deux premières sont responsables de la définition de la stratégie et de la programmation des aides européennes. EuropeAid est responsable de la mise en œuvre des instruments d'aide (depuis l'identification des projets et des programmes jusqu'à leur évaluation *ex post*) pour l'ensemble des pays tiers.

L'évaluation des interventions de l'aide européenne est assurée à deux niveaux: a) l'évaluation des projets qui relève de la responsabilité des délégations<sup>12</sup> qui les mettent en œuvre; b) les évaluations de synthèse ainsi que le suivi externe de projets qui sont réalisés sous le contrôle d'une unité d'évaluation centrale. Ces évaluations sont confiées à des évaluateurs externes.

L'unité d'évaluation, commune aux trois DG évoquées ci-dessus, rend compte directement au Commissaire en charge des Relations extérieures. Elle fournit un appui méthodologique aux délégations et aux responsables des évaluations de projets. Pour ce faire, l'unité développe actuellement une série d'outils de collecte d'information et d'analyse, présentés sous la forme de lignes directrices méthodologiques qui seront accessibles sur le site Internet de la Commission européenne à la fin de 2005. Cet effort de formalisation des méthodes prend toute son importance dans un contexte où la Commission européenne déconcentre une série de responsabilités, y compris les évaluations, vers ses délégations dans les pays tiers.

L'unité d'évaluation d'EuropeAid bénéficie elle-même de l'appui méthodologique de l'unité correspondante de la DG Budget. Cette dernière est

considérée comme la tête d'un réseau d'évaluation couvrant l'ensemble des DG de la Commission.

Les déclarations politiques du début des années 2000 ont donc été largement traduites dans les faits, à tout le moins en ce qui concerne le montage institutionnel. L'évaluation de projets est toutefois restée le maillon faible, partiellement compensée par un système de suivi coordonné par l'unité d'évaluation.

Ce système, baptisé le ROM<sup>13</sup>, consiste en de brèves visites sur le site des projets par des consultants extérieurs qui remplissent des fiches de notation standardisées et où ils fournissent un certain nombre d'informations factuelles. Ils doivent également donner leur appréciation de l'efficacité, de l'efficience, de l'impact, de la pertinence et de la durabilité probable des projets et des programmes. Leurs travaux reposent sur l'analyse de documents et sur des entretiens avec des représentants des principales catégories de participants aux projets, y compris les bénéficiaires finaux. Les rapports, les avis des gestionnaires et les documents fondamentaux des projets sont consignés dans une base de données centrale. Le suivi porte sur les projets en cours (mis en œuvre depuis au moins six mois et prévus pour une période équivalente) et dotés d'un budget minimal d'environ 1 million d'euros. En 2004, 1 100 projets d'une valeur totale d'environ 7,6 milliards d'euros ont été suivis dans 117 pays répartis dans les 6 régions de coopération.

Ce système constitue un progrès important par rapport au passé en termes de gestion de l'information sur les projets. On peut toutefois se demander s'il est indispensable que les consultants, compte tenu des moyens dont ils disposent pour chaque intervention, dépassent leur rôle de moniteur chargé de suivre un certain nombre de paramètres pour se lancer dans une appréciation selon des critères de type évaluatif tels que ceux cités plus tôt.

<sup>12</sup> Représentations de la Commission européenne dans les pays tiers.

<sup>13</sup> *Results Oriented Monitoring*.

## Une approche adaptée à des objets complexes

Les évaluations dont il est question ici portent essentiellement sur les stratégies d'aide adoptées par la Commission européenne, abordées soit sur le plan sectoriel ou thématique, soit au niveau d'un pays ou d'une région<sup>14</sup>. Parmi les évaluations thématiques déjà réalisées, on citera les interventions en appui au développement des transports, à la protection de l'environnement, au développement du secteur privé, à la sécurité alimentaire, les programmes de microréalisations, les programmes d'appui au développement des capacités commerciales, etc.

Les champs couverts par ces évaluations sont étendus et hétérogènes. La période sous revue est souvent de dix ans. Pour ce qui est d'un pays ou d'une région, les évaluations couvrent la stratégie de coopération de la Commission européenne et sa mise en œuvre. Ceci implique la prise en compte de toutes les interventions communautaires et de tous les instruments utilisés dans les différents secteurs d'intervention dans ce pays ou cette région. Sur le plan sectoriel ou thématique, les évaluations se concentrent sur la stratégie et la mise en œuvre de la coopération dans un secteur spécifique, mais portent, en général, sur tous les pays et régions recevant de l'aide européenne.

Les sources d'hétérogénéité des champs d'évaluation sont donc de plusieurs ordres :

- Temporel d'abord : dix ans, c'est très long. Une telle période chevauche toujours les années 1998-2002 au cours desquelles les principales réformes de l'approche de l'aide européenne ont été formulées dans la plupart des secteurs. En outre, de nombreux pays partenaires ont connu des crises politiques ou naturelles majeures durant cette période.

- Spatial ensuite : l'appui au développement du secteur privé au Vietnam, en Zambie ou au Maroc pose des problèmes assez différents. L'évaluation d'une politique sectorielle appliquée dans des contextes aussi différents peut conduire à des conclusions très contrastées selon les régions.
- Sectoriel enfin : dans un pays donné, porter un jugement sur un spectre d'activité allant de la construction de routes à la promotion des droits de l'homme n'est pas toujours facile.

Ces difficultés ne condamnent toutefois pas l'idée de pratiquer des évaluations à ces niveaux. En effet, la Commission européenne a amélioré la cohérence de son action en affichant ses politiques d'intervention à l'échelle de l'ensemble des pays tiers et pour chacun d'entre eux. Il est donc logique qu'elle se pose périodiquement la question de la pertinence, de l'efficacité ou de l'efficience de ces politiques. L'important est d'admettre l'hétérogénéité des champs couverts et d'en déduire une approche évaluative qui aura le souci d'éviter la généralisation de constats trop spécifiques.

## Les priorités politiques exprimées dans les critères d'évaluation

Dans tous les cas, l'analyse est basée sur les critères d'évaluation proposés par le CAD<sup>15</sup>, à savoir : la pertinence, l'efficacité, l'impact, la durabilité et l'efficience. La notion de pertinence est souvent entendue comme la cohérence entre l'objectif spécifique<sup>16</sup> de l'objet d'évaluation (la stratégie de coopération dans un secteur ou la stratégie de coopération avec un pays) et les objectifs plus généraux<sup>17</sup> de la coopération européenne avec les pays tiers. L'analyse d'efficacité est généralement centrée sur la mesure dans laquelle les objectifs opérationnels et dans certains cas l'objectif spécifique de l'objet d'évaluation sont atteints.

<sup>14</sup> La notion de région étant entendue ici comme un groupement de plusieurs pays voisins bénéficiant d'accords de coopération avec la Communauté européenne (l'Union Économique et Monétaire Ouest Africaine, Mercosur, les pays méditerranéens).

<sup>15</sup> Development Assistance Committee (1991).

<sup>16</sup> Objectif spécifique : objectif que l'intervention doit atteindre.

<sup>17</sup> Objectif général : objectif politique d'un ordre supérieur auquel l'intervention doit contribuer.

À ces critères classiques s'ajoutent des priorités plus propres à l'aide européenne telles que la cohérence entre les politiques de l'Union qui peuvent affecter les pays en développement (politiques commerciale, agricole ou autre) et les objectifs de la politique de développement et la coordination des interventions entre la Communauté et les autres bailleurs de fonds.

Les traités indiquent que «L'Union veille, en particulier, à la cohérence de l'ensemble de son action extérieure dans le cadre de ses politiques en matière de relations extérieures, de sécurité, d'économie et de développement. Le Conseil et la Commission ont la responsabilité d'assurer cette cohérence et coopèrent à cet effet»<sup>18</sup>. On lit en outre que «La Communauté et les États membres coordonnent leurs politiques en matière de coopération au développement et se concertent sur leurs programmes d'aide, y compris dans les organisations internationales et lors des conférences internationales»<sup>19</sup>.

La politique de développement de la Communauté européenne reprend ces exigences et rappelle que «*La réalisation de cet objectif [de développement] suppose une cohérence accrue entre l'ensemble de nos politiques ayant un impact sur les pays en développement. Une telle cohérence dépend également d'une meilleure coordination et d'une complémentarité avec les actions des États membres*»<sup>20</sup>.

Finalement, les évaluateurs doivent également apprécier la mesure dans laquelle les interventions financées par la Commission contribuent à promouvoir des valeurs telles que les droits de l'homme, la bonne gouvernance, l'égalité entre les

sexes ou la protection de l'environnement<sup>21</sup>. Il s'agit de valeurs définies comme importantes par les représentants de ce que nous avons appelé plus haut les «citoyens-contribuables». La Commission doit donc vérifier si son aide contribue à la promotion de ces valeurs.

### Des interventions évaluées en référence à la politique affichée

La politique européenne d'aide au développement se décline par secteur ou thème d'intervention tels que le développement rural, les transports, l'appui au secteur privé, la gestion de l'eau<sup>22</sup>, etc. Ces documents de politique sectorielle servent de référence lors de la conception et de la négociation d'interventions avec les pays bénéficiaires et les autres bailleurs de fonds. Ils constituent également la référence principale des évaluations sectorielles. Ils sont de plus en plus structurés sur une base standardisée qui impose à leurs auteurs d'afficher relativement explicitement leurs objectifs.

Dans le cas de la stratégie de coopération avec un pays, le texte de référence est intitulé Document de stratégie pays (DSP)<sup>23</sup>. Ce document est produit tous les cinq ans par la délégation de la Commission européenne et fait l'objet d'un accord avec le gouvernement bénéficiaire. Une revue de l'exécution de ces accords est prévue à mi-parcours, au moins pour les pays d'ACP. En outre, un rapport annuel d'exécution est produit conjointement avec le gouvernement bénéficiaire pour analyser le progrès réalisé dans la mise en œuvre du programme de coopération.

<sup>18</sup> Union Européenne (2002b), Titre I, Article 3.

<sup>19</sup> Union Européenne (2002a), Titre XX, Article 180.

<sup>20</sup> Commission européenne (2000a), p. 4.

<sup>21</sup> Ces valeurs sont curieusement appelées *crosscutting issues* plus ou moins bien traduit par «aspects transversaux».

<sup>22</sup> Voir respectivement Commission européenne (2002b, 2000b, 2003, 2002a).

<sup>23</sup> Jusqu'à 2002, ces documents ne comprenaient que des descriptions et des analyses de la situation du pays. La programmation des interventions, établie en accord avec le pays bénéficiaire, était consignée dans un document séparé appelé Programme Indicatif National (PIN). Actuellement, les deux dimensions sont contenues dans un seul document.

Les procédures actuelles sont relativement strictes : la première tâche des évaluateurs est de présenter la logique d'intervention, autrement dit l'arbre des objectifs caractérisant l'action européenne à partir des documents de politique cités ci-dessus. La logique d'intervention ainsi présentée constitue la principale base de référence pour juger de la pertinence et de la cohérence interne de la politique. On s'y référera également pour juger de l'efficacité de la politique et pour établir les questions d'évaluation.

### **Une analyse guidée par des questions d'évaluation**

Les questions d'évaluation servent à mieux focaliser le travail de l'évaluateur sur un nombre limité de sujets. En général, le nombre de questions est limité à dix, ce qui est beaucoup, mais traduit les débats parfois animés qui ont précédé la commande de l'évaluation. On essaie autant que possible de les formuler en fonction des critères d'évaluation. Les questions correspondant à la pertinence, à l'efficacité, à l'impact et à la durabilité sont ancrées dans la logique d'intervention. Leur formulation dépend donc de la nature du secteur étudié et des objectifs poursuivis. Par contre, les questions correspondant aux autres critères tels que l'efficacité, la coordination et les thèmes transversaux sont généralement moins dépendantes du programme évalué. Les questions évaluatives sont formulées soit par la Commission lors de la commande de l'évaluation, auquel cas elles apparaissent dans les termes de référence (cahier des charges), soit par les évaluateurs dans une première phase de leur démarche (phase de structuration).

### **La Commission seule aux commandes**

Les trois partenaires habituels de ce type d'évaluation sont l'unité chargée de l'évaluation de l'aide européenne (unité d'évaluation), le groupe de référence et l'équipe d'évaluation :

- *L'unité d'évaluation*, équipe spécialisée maîtrisant les méthodes, est chargée de conduire l'opération. Pour chaque évaluation, elle constitue le groupe de référence décrit ci-dessous, formule

les termes de référence, sélectionne les évaluateurs, assure le suivi de leur travail, approuve les rapports à chaque étape du processus et en assure la restitution interne et la publicité.

- *Le groupe de référence*, exclusivement composé de fonctionnaires de la Commission européenne est en charge des interventions concernées par l'évaluation. Dans le cas des évaluations de programmes nationaux, les délégations européennes dans ces pays et des représentants nationaux sont également présents. Ce groupe commente les rapports à différents stades de l'évaluation. Ses membres peuvent également fournir un relais facilitant l'accès aux services concernés par l'étude. Le groupe de référence est présidé par un représentant de l'unité d'évaluation.
- *Les évaluateurs*, équipe composée de consultants externes à la Commission.

Il s'agit donc d'évaluations externes, mais dont le contrôle est exclusivement entre les mains de la Commission européenne. Ce contrôle n'implique aucune représentation formelle ni des bénéficiaires ni même des acteurs non membres de la Commission. Ceux-ci sont consultés, autant que possible, vu l'étendue des champs couverts, par les évaluateurs, mais n'ont aucune influence directe sur la démarche. Ce type d'évaluation pose des problèmes méthodologiques particulièrement complexes, notamment en raison de l'ampleur et de l'hétérogénéité des champs couverts. La conception de l'échantillon de pays et d'interventions à observer et le choix des méthodes de collecte d'information et d'analyse reposent sur des bases méthodologiques encore en pleine évolution.

Le cadre institutionnel présenté ci-dessus, impliquant un contrôle quasi exclusif de l'évaluation par la Commission européenne, par ailleurs principal concepteur et gestionnaire des politiques évaluées, ne serait clairement pas adapté à une démarche évaluative « sommative » (destinée à rendre compte au citoyen). Ce n'est pas l'objet principal des évaluations décrites ici. En effet, en référence aux différents rôles de l'évaluation présentés ci-dessus, nous sommes dans le dernier cas, décrivant les évaluations formatives réalisées entre



acteurs de l'aide afin d'améliorer la qualité de leurs décisions. Il s'agit d'une démarche dont la méthodologie n'est pas encore stabilisée qui vise à permettre à la Commission de tirer les leçons de son expérience et à alimenter, sur cette base, les réflexions qui ont accompagné la réforme profonde de la conception de l'aide au développement qui a caractérisé la décennie 1995-2005.

La logique d'intervention, les questions d'évaluation, l'approche prévue par les évaluateurs pour y répondre, les sources utilisées et, plus tard, les constats, sont proposés au groupe de référence qui les commente. L'unité d'évaluation valide ensuite chaque étape du processus, confirmant que les termes de référence sont exécutés selon des standards professionnels acceptables. Que penser de ces validations successives? Ces évaluations complexes qui engagent des équipes parfois nombreuses pour 12 à 18 mois, comportent des risques pour toutes les parties. L'évaluateur, tel un grimpeur qui escalade une paroi rocheuse, doit y planter régulièrement des pitons et y passer sa corde pour éviter de retomber trop bas en cas de chute. La validation à différentes étapes de l'évaluation joue ce rôle. La solidité des validations dépend largement de l'unité d'évaluation qui s'assure à chaque stade que l'évaluateur travaille selon les standards professionnels les plus stricts, réduisant par là son propre risque et celui des autres parties prenantes. Mais une fois cette étape validée, l'unité d'évaluation « protège » l'évaluateur contre la remise en cause ultérieure de cette étape par une des parties concernées. Les validations successives assurent que la construction progressive de la confiance entre les parties prenantes, indispensable à la production d'une évaluation crédible, s'établit sur des bases solides. Elles permettent à l'évaluateur de continuer à prendre des risques jusqu'à la fin du processus et donc accroissent à la fois son indépendance et la crédibilité de l'évaluation.

Si la conduite de l'évaluation reste assez confinée au sein des services de la Commission, un effort sensiblement accru est consenti depuis peu pour en assurer la diffusion et le suivi des résultats.

### Un effort de publicité et de rétroaction

Les standards d'évaluation de la Commission européenne<sup>24</sup> requièrent l'établissement de mécanismes de diffusion, de rétroaction et de suivi appropriés pour toutes les évaluations. Cinq canaux de dissémination sont établis :

- Les rapports sont d'abord envoyés aux commissaires et aux directeurs généraux, ensuite aux services directement concernés par l'évaluation ainsi qu'au réseau d'évaluation du CAD. Les rapports sont également transmis sur demande au Parlement européen et à la Cour des comptes;
- Les rapports d'évaluation sont ensuite mis à la disposition du public sur le site Internet de la Commission européenne<sup>25</sup>. Une fiche contradictoire (voir plus bas) élaborée par les services concernés y est jointe;
- Après la publication du rapport sur Internet, un court article contenant l'objet, le champ couvert et les conclusions de l'évaluation est publié dans l'intranet des DG concernées par l'aide au développement;
- Le même article est publié dans la *Newsletter* interne d'EuropeAid. En outre, l'unité d'évaluation produit tous les six mois une « *Newsletter* de rétroaction » avec des articles sur toutes les évaluations récemment produites;
- Une fois que le rapport d'évaluation a été distribué, un atelier peut être organisé à Bruxelles ou dans le pays dont le programme est évalué pour encourager la discussion autour des constats et des recommandations de l'évaluation. Des représentants des États membres, des pays bénéficiaires, des autres bailleurs de fonds et des ONG sont notamment invités selon le sujet.

<sup>24</sup> Commission européenne (2002c).

<sup>25</sup> Commission européenne, <http://europa.eu.int/comm/europeaid/evaluation/index.htm>

La rétroaction est aussi importante que la dissémination. Elle contribue à la prise en compte des résultats des évaluations et permet à l'unité d'évaluation de s'informer sur la manière dont les évaluations sont utilisées par les services concernés. L'outil de rétroaction le plus important est la fiche contradictoire. Ce document doit être rempli pour toutes les évaluations terminées. Il comporte trois colonnes : a) la première liste les recommandations de l'évaluation; b) la deuxième présente les réponses des services concernés par ces recommandations; c) la troisième, remplie un an plus tard, présente les actions entreprises par les services. La fiche contradictoire est publiée sur le site Web de la Commission européenne avec le rapport d'évaluation. L'unité d'évaluation a la responsabilité de demander aux services concernés de remplir la fiche contradictoire immédiatement après la finalisation de l'évaluation et un an plus tard.

Une visite du site Web indiqué précédemment permet de vérifier que les rapports sont effectivement publiés et que les fiches contradictoires sont établies. Par contre, on trouve encore peu de signes de suivi d'exécution des recommandations après un an.

L'effort consenti pour systématiser tant la valorisation interne que la publicité des évaluations est incontestable et donne une crédibilité certaine à l'ensemble de la démarche.

### ■ DÉFIS À RELEVER ET PISTES À EXPLORER

En quinze ans, le chemin parcouru est énorme, mais les défis encore à relever pour améliorer la qualité et pouvoir systématiser le recours aux évaluations de synthèse le sont tout autant. On peut constater notamment :

- le problème d'« imputabilité », ou à qui doit-on rendre compte,
- l'évolution de la nature de l'aide européenne,

- la mauvaise qualité des données et l'absence de normes,
- la dispersion géographique des interventions et donc le coût de leur évaluation,
- l'hétérogénéité des champs des grandes évaluations thématiques.

À qui rendre compte ? On l'a vu, le bénéficiaire a peu de pouvoir. Il n'a aucun pouvoir « revendicatif » tel que celui qu'ont obtenu les bénéficiaires des solidarités intranationales, porteurs de mouvements sociaux qui ont réussi à s'imposer et à obtenir les transferts sociaux que nous connaissons. Dans le contexte international, il est donc essentiel de rendre compte à celui que nous avons appelé le citoyen-contribuable ou à ses représentants contrôlant les cordons de la bourse. Toutefois, un système d'évaluation strictement justifié par la volonté de maintenir, voire d'augmenter le flux d'aide, risque de devenir rapidement « cosmétique », plus enclin à justifier qu'à juger<sup>26</sup>. Certains considèrent que c'est déjà largement le cas. Le défi est donc bien d'axer l'imputabilité sur le contribuable, tout en s'assurant que l'évaluation reste critique. Pour éviter les dérives complaisantes, il est fondamental d'intégrer autant que possible l'expression des bénéficiaires dans ces évaluations et de leur permettre d'y réagir.

Le défi à relever par l'évaluation et la gestion des flux d'information qui en sont le corollaire est donc de faciliter l'expression du bénéficiaire et de le mettre en communication avec le contribuable pour qu'ensemble ils puissent assurer un meilleur contrôle de la chaîne d'acteurs. Il faut pour cela préciser les zones de convergences d'intérêt entre les uns et les autres.

Les ONG qui investissent beaucoup dans la médiation de leur action ont compris l'enjeu de cette communication. Elles cherchent à motiver le contribuable en donnant un visage au bénéficiaire et en lui accordant un espace de communication,

<sup>26</sup> Martens et al. (2002).

même si elles en contrôlent le contenu de manière parfois discutable. L'évaluation de l'aide doit être plus indépendante et plus critique, mais doit jouer sur les mêmes registres. Elle doit donc être réalisée plus largement par les gouvernements et surtout par les populations bénéficiaires, mais pas exclusivement, et même pas essentiellement pour eux, si l'on veut que les acteurs de l'aide soient réellement soumis à un contrôle externe et gardent leur crédibilité.

Comment adapter l'évaluation à l'évolution des approches? La Commission européenne est engagée dans un processus de déconcentration et de décentralisation de la gestion de son aide. La déconcentration consiste à donner plus de responsabilités aux délégations et la décentralisation à responsabiliser plus directement les autorités locales. De plus en plus, l'aide prend la forme d'appuis budgétaires finançant des politiques macroéconomiques ou sectorielles, ce qui implique plus de leadership du gouvernement bénéficiaire, plus de coordination entre bailleurs de fonds, moins de possibilités pour chaque bailleur d'affecter ses fonds à des opérations précises, moins de contrôle des ressources, mais plus de suivi des résultats. Parallèlement, et non sans certaines contradictions, la tendance est également à la prise en charge de missions de service public par des organisations privées issues de la population ou des ONG internationales. Où faut-il, dans un tel contexte, situer la fonction d'évaluation? Au niveau central? Au niveau des délégations? Comment procéder?

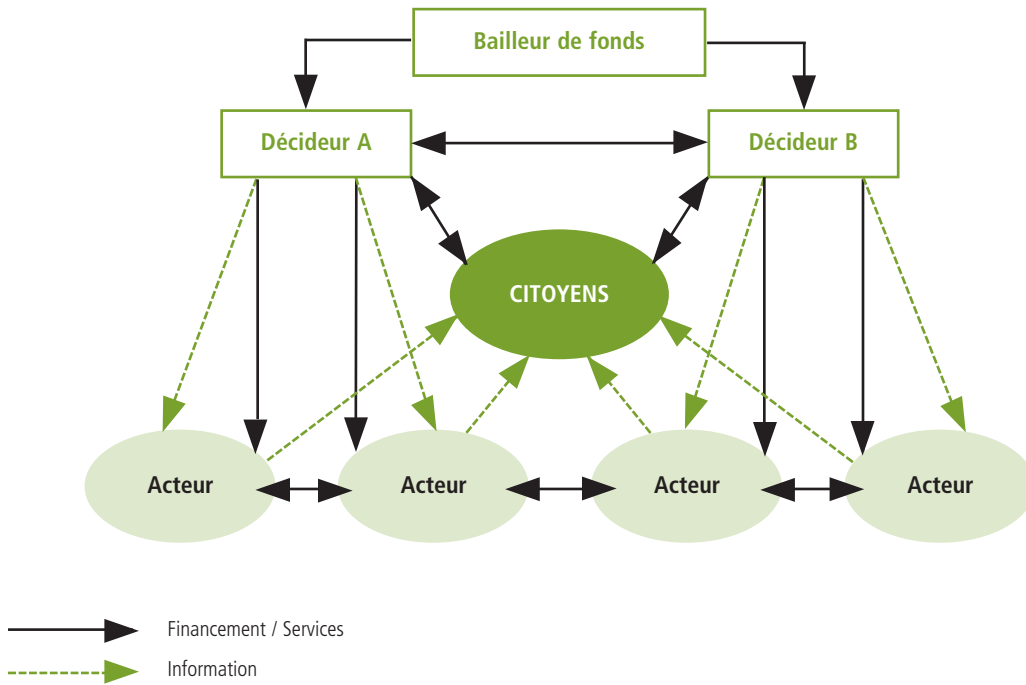
Comment traiter le problème de la qualité des données? Voici peu, la pression sur les acteurs de l'aide internationale, pour qu'ils assurent un suivi systématique de leurs interventions et surtout des effets de celles-ci, était très faible. Il est souvent difficile, par exemple, de reconstituer même l'inventaire des interventions réalisées dans un secteur donné au cours de ces dix dernières années. Plus récemment, au contraire, les nécessités de la coordination multi bailleurs, le dialogue politique avec les gouvernements et la culture de gestion par objectif ont donné lieu à l'identification d'une plé-

thore d'indicateurs pour lesquels nous ne disposons que de séries courtes et souvent discontinues. Le volume de données s'accroît incontestablement, mais comment juger sur cette base? Sur quelles normes s'appuyer pour convertir ces indicateurs en critères de jugement, crédibles et acceptables?

Les évaluations de programmes nationaux comme les évaluations thématiques souffrent, on l'a déjà évoqué, de la grande hétérogénéité du champ qu'elles doivent couvrir. La solution logique à ce problème est soit de diviser le champ en périodes (espaces) plus homogènes, soit d'augmenter la taille des échantillons d'observation pour pouvoir analyser et conclure sur une base de constats suffisamment nuancés. Tout cela coûte cher et, même si les budgets d'évaluation ont augmenté ces dernières années, ils constituent encore une contrainte sévère. La Commission européenne réalise actuellement environ 5 évaluations de programmes pays par an alors qu'elle intervient dans plus de 120 pays, ce qui implique pour chaque pays une fréquence théorique d'évaluation de une tous les 24 ans. Nous sommes loin des six ans évoqués précédemment qui impliqueraient un nombre annuel d'évaluations quatre fois supérieur.

Comment faire face à ces divers défis à un coût acceptable? La Commission européenne est, nous l'avons vu, un bailleur de fonds important qui intervient dans de nombreux pays dans lesquels elle applique une stratégie d'intervention, finance des programmes, développe des projets. La Commission dispose d'assez d'information pour tenter de regrouper ces stratégies, ces interventions, en grappes (clusters) suffisamment homogènes pour permettre certaines comparaisons. Que l'on parle de stratégies pays ou d'interventions dans des secteurs tels que l'éducation ou les transports, le schéma est toujours le même (figure 3).

FIGURE 3 : MODÈLE DE RÉSEAU D'ÉVALUATION



Le bailleur de fonds finance des décideurs (gouvernements, autorités locales) qui eux-mêmes délèguent à des acteurs (entreprises, ONG, associations locales...) l'exécution des interventions d'aide au développement. L'évaluation que nous avons évoquée au cours de cet article est une analyse ponctuelle, généralement commandée par le bailleur de fonds qui porte sur la relation entre ce dernier et le décideur pour exécuter un programme. Le comportement des acteurs de terrain alimente la réflexion d'ensemble, mais ceux-ci retirent peu de leçons d'une évaluation qui porte sur des enjeux plus larges.

Chaque évaluation est traitée indépendamment, ce qui pose les problèmes évoqués ci-dessus et en particulier le problème des normes : par rapport à quel niveau de performance peut-on considérer que telle intervention, telle stratégie-pays est efficace, efficiente, etc.? Le jugement de l'évaluateur s'appuie sur les objectifs affichés ou sur des nor-

mes plus ou moins explicites basées sur sa connaissance d'expériences comparables.

Peut-on systématiser ces comparaisons en stimulant une pratique d'étalonnage régulier? La mise en communication de délégations chargées de mettre en œuvre des stratégies d'aide dans des pays considérés comme comparables leur permettrait de réaliser ensemble l'évaluation de l'action européenne dans ces pays. Cette approche, conduite avec un support professionnel, pourrait apporter plus qu'une série d'évaluations indépendantes les unes des autres et amorcer la constitution sur Internet de plateformes d'échanges d'information et d'analyses entre décideurs ou acteurs appartenant aux mêmes groupes. Dans le même esprit, peut-on mettre systématiquement en communication les responsables de politiques sectorielles telles que l'éducation, la santé ou les transports de différents pays?

L'évaluateur ne serait plus chargé d'effectuer les évaluations les unes après les autres, mais il aurait plutôt la responsabilité de réaliser les analyses nécessaires à la constitution des groupes, d'établir les contacts entre décideurs et entre acteurs et ensuite d'animer les réseaux en proposant questions, critères de jugement et indicateurs, en collectant, redistribuant et analysant les données fournies et en conduisant progressivement ces réseaux vers une forme de *peer review* permanente conduite avec professionnalisme. L'évaluateur pourrait également s'assurer que ces réseaux soient ouverts et communiquent activement, par des canaux appropriés, avec les représentants des citoyens tant bénéficiaires que contribuables.

Bien entendu, il faudrait motiver les parties prenantes à jouer le jeu de l'échange d'information. La Commission européenne, comme d'ailleurs les autres bailleurs de fonds, dispose des moyens nécessaires pour encourager la participation. Tant la carotte que le bâton sont de même nature que ceux utilisés pour encourager l'ouverture à l'évaluation : capacité d'expression et analyse de l'expérience d'un côté, condition à respecter pour se voir renouveler le financement de l'autre.

Ces plateformes coopératives de suivi et d'évaluation comparatives pourraient constituer de puissants instruments de rapprochement contribuant à refermer la « boucle de rétroaction » évoquée au début de cet article. Le processus de mondialisation que nous observons ne conduit pas à la mondialisation de l'espace politique. Les contribuables et les bénéficiaires de l'aide au développement appartiendront encore longtemps à des espaces politiques distincts. Tant que cette boucle restera aussi ouverte, l'aide au développement restera une démarche technocratique peu susceptible de motiver l'accroissement considérable des transferts sociaux nécessaires pour réduire la pauvreté. L'évaluation et les démarches qui pourraient en découler garderont donc un rôle important à jouer.

## BIBLIOGRAPHIE

- ADE (1995). *Evaluation of agricultural development projects funded during the Lomé I, II & III cycles*, Brussels.
- Commission européenne (2000a). *Communication de la Commission au Conseil et au Parlement européen, La politique de développement de la Communauté européenne*, COM (2000) 212, Bruxelles.
- Commission européenne (2000b). *Communication de la Commission au Conseil et au Parlement européen, La promotion de transports durables dans la coopération au développement*, COM (2000) 422, Bruxelles.
- Commission européenne (2000c). *Communication à la Commission de Mme Schreyer en accord avec M. Kinnock et le Président, Mettre l'accent sur les résultats : renforcer l'évaluation des activités de la Commission*, SEC (2000) 1051, Bruxelles.
- Commission européenne (2000d). *Statement by the Council and the Commission: The European Community's Development Policy*, Brussels.
- Commission européenne (2002a). *Communication de la Commission au Conseil et au Parlement européen, La Gestion de l'eau dans les pays en développement: Politique et priorités de la coopération au développement de l'UE*, COM (2002) 132, Bruxelles.
- Commission européenne (2002b). *Communication de la Commission, Lutte contre la pauvreté rurale: Politique et approche de la Communauté européenne à l'égard du développement rural et de la gestion durable des ressources naturelles dans les pays en développement*, COM (2002) 429, Bruxelles.
- Commission européenne (2002c). *Communication à la Commission de M. le Président et de Mme Schreyer, Standards et bonnes pratiques d'évaluation*, COM (2002) 5267, Bruxelles.
- Commission européenne (2003). *Communication de la Commission au Conseil et au Parlement européen, Coopération de la Communauté européenne avec les pays tiers: Comment la Commission envisage de soutenir, à l'avenir, le développement des entreprises*, COM (2003) 267, Bruxelles.
- Commission européenne (2005). *Rapport annuel 2005 sur la politique de développement de la Communauté européenne et la mise en œuvre de l'aide extérieure en 2004*, Bruxelles.
- Development Assistance Committee (1991). *Principles for Evaluation of Development Assistance, Organisation for Economic Cooperation and Development*, Paris.
- Dollar, D. et L. Pritchett (1998). *Assessing Aid, What Works, What Doesn't, and Why?*, Oxford University Press, Oxford.
- Martens, B. et al. (2002). *The Institutional Economics of Foreign Aid*, Cambridge University Press, London.
- Tecnecon (1995). *Evaluation of EDF transport sector projects and programmes*, Brussels.

Union Européenne (2002a). *Version Consolidée du Traité Instituant la Communauté européenne, Journal Officiel des Communautés Européennes*, C325/33, Bruxelles.

Union Européenne (2002b). *Version Consolidée du Traité sur l'Union européenne, Journal Officiel des Communautés Européennes*, C325/33, Bruxelles.

## Prochain numéro : automne 2006

### **L'ÉVALUATION DE LA GOUVERNANCE DANS LES « PETITES SOCIÉTÉS »**

Nous consacrons le prochain numéro de *Télescope* (automne 2006) à une analyse comparée de la gouvernance dans les sociétés de petite taille. Ici, l'unité de mesure est bien sûr le nombre d'habitants ou la superficie du territoire. Une faible population ou un espace réduit n'empêchent pas ces pays d'être grands dans l'histoire comme dans le présent. Un présent conjugué d'ailleurs le plus souvent en termes de stabilité politique, de prospérité économique et de consensus social.

Norvège, Suisse, Danemark, Irlande, Suède, Belgique, autant de *success stories*, autant de pays dont on essaie de comprendre, et parfois même de reproduire, les recettes. La mondialisation contribuant à faire de tous les pays des « petits pays », la gouvernance de ces sociétés en équilibre prend valeur d'exemple. Mais quelle est donc justement la clef de leur réussite ?

Invités par *Télescope* à répondre à cette question, les experts livrent leur version de l'évolution récente de la gouvernance de ces pays, c'est-à-dire de la représentation de l'État à sa société. Partout l'État « diminue ». Devenu stratège, il se dégage de l'opérationnel, décentralise, privatise, conclut des partenariats et renouvelle ses modes d'intervention et de délivrance des services publics. Pour certains, on assiste à la fin du *Big Government*. D'autres n'y voient qu'un simple transfert du *Big Government* du niveau central vers les niveaux régionaux ou locaux.

Accompagnateur ou régulateur et de moins en moins producteur ou opérateur, l'État se repositionne dans sa société. Cette transformation de la relation de l'État avec la société civile est une constante des témoignages recueillis. Ici, on parle de nouveau contrat social, là, de *Negotiated Economy*, ailleurs encore de participation citoyenne ou de « démocratie à temps plein ».

C'est peut-être à la conjonction du recentrage de l'État et du renforcement d'une démocratie participative venant s'ajouter à la démocratie représentative que réside le « secret » du succès des « petits pays ».



*Télescope* est une revue d'analyse comparée portant sur des problématiques spécifiques intéressant les administrations publiques. Elle est publiée quatre fois par année. Sa production est coordonnée par l'équipe du département des publications de L'Observatoire de l'administration publique. Le choix des thèmes et des textes de *Télescope* font l'objet d'une réflexion collégiale de la part des membres de L'Observatoire. Avant publication, tout article est soumis à un comité composé d'universitaires qui évalue son acceptabilité. En cas de controverse sur un article ou sur une partie d'un article, l'auteur est avisé. La révision linguistique est obligatoire. Elle est assurée par les services spécialisés de l'ENAP. La reproduction totale ou partielle de *Télescope* est autorisée avec mention obligatoire de la source. Les professeurs d'établissements d'enseignement ne sont pas tenus de demander une autorisation pour distribuer des photocopies. *Télescope* bénéficie du soutien financier de L'ENAP du Québec et du gouvernement du Québec.

L'OBSERVATOIRE DE L'ADMINISTRATION PUBLIQUE  
TÉLESCOPE

DÉPÔT LÉGAL  
BIBLIOTHÈQUE NATIONALE DU QUÉBEC  
BIBLIOTHÈQUE NATIONALE DU CANADA  
ISSN 1203-3294



L'OBSERVATOIRE DE  
L'ADMINISTRATION  
PUBLIQUE  
E N A P