

## INSTITUTIONNALISATION DE L'ÉVALUATION ET NOUVELLE GESTION PUBLIQUE : UN ÉTAT DES LIEUX COMPARATIF

Frédéric VARONE et Steve JACOB

*Cet article compare le degré d'institutionnalisation de l'évaluation des politiques publiques dans dix-huit démocraties. Trois catégories de pays sont ainsi identifiées (pionniers, intermédiaires et retardataires). Afin d'interpréter ces différences, six facteurs explicatifs sont discutés : la pratique de la Nouvelle Gestion Publique, le niveau des déficits budgétaires, les pressions externes, le type majoritaire ou consensuel ainsi que la structure fédéraliste ou unitaire du régime démocratique et l'indépendance de la Cour des comptes. Si chacune de ces hypothèses semble plausible pour comprendre la situation de l'évaluation dans un pays donné, force est de constater qu'aucune ne résiste à une analyse portant sur l'ensemble des systèmes politiques retenus ici. En conclusion, une étude approfondie du rôle des réformateurs du secteur public s'avère indispensable.*

### Évaluer les réformes, réformer par l'évaluation

Plusieurs publications scientifiques et rapports d'organismes internationaux se penchent sur l'évaluation de la Nouvelle Gestion Publique (NGP) en tant que réforme managériale du secteur public. Cette entreprise critique nous semble des plus louables, ne serait-ce que pour dépasser les différends, parfois stériles, sur l'acceptabilité ou non des postulats, principes d'action et outils opérationnels des actuelles mutations politico-administratives. Le recours à l'évaluation de la "NGP au concret" ou de la "NGP en action" devrait représenter un passage obligé, afin de porter des jugements de valeur solides, basés sur des constats empiriques dûment étayés. Par conséquent, nous souscrivons avec empressement au constat de C. Pollitt<sup>1</sup> qui exhorte les thuriféraires comme les plus farouches adversaires de la NGP à s'engager dans une appréciation systématique et transparente des expérimenta-

---

1. POLLITT C., "Justification by Works or by Faith ? Evaluating the New Public Management", *Evaluation. The International Journal of Theory, Research and Practice*, 1995, vol. 1, n° 2, pp. 133-154.

tions en chantier. Et ce plutôt que de se limiter à encenser, à des fins stratégiques de légitimation *ex post* de la NGP, quelques bonnes pratiques en la matière ou, *a contrario*, de jeter l'opprobre sur l'ensemble de ces démarches réformatrices, dont les transcriptions pratiques prennent des formes plurielles au demeurant, en soulignant des échecs retentissants dans tel ou tel secteur d'action publique et pays. De fait, les rares évaluations rigoureuses de l'application – encore partielle et provisoire dans plusieurs systèmes politico-administratifs – des concepts et instruments de la NGP dressent généralement des bilans contrastés qui, in fine, invitent tant à renforcer certaines pratiques de NGP ayant fait leur preuve que, concomitamment, à en rejeter d'autres dont l'inefficacité est patente.

Par contre, force est de noter que les chercheurs et plus encore les décideurs politico-administratifs n'expriment qu'un intérêt assez limité pour comprendre et expliquer, en tous les cas sur une base comparative, en quoi – dans des contextes institutionnels et constellations d'acteurs particuliers – la NGP facilite ou au contraire inhibe l'utilisation de l'évaluation des politiques publiques<sup>2</sup> pour moderniser le secteur public ou, du moins, accompagner les réformes de l'État. Cet article participe donc à la tentative d'identifier le rôle concrètement assigné à la pratique évaluative comme (éventuel) élément (moteur) des réformes prônées dans la conduite de l'action publique, en tenant compte d'autres dimensions institutionnelles et jeux d'acteurs bien entendu. De ce point de vue, plusieurs questions mériteraient d'être abordées. À titre exemplatif, nous recensons les interrogations suivantes, dont les enjeux théoriques et la portée pratique pour la transformation du secteur public nous paraissent pertinents : comment l'évaluation s'inscrit-elle dans la philosophie et la stratégie d'une meilleure gestion de la performance publique qui, par essence, présuppose une mesure et une appréciation des effets induits par les politiques publiques et pas uniquement de la rationalisation interne des services administratifs et des prestations publiques ? L'évaluation est-elle favorisée par certains types de réformes de NGP ou, au contraire, abandonnée suite à sa mise en concurrence avec d'autres outils de management public et de pilotage des politiques publiques comme les multiples tableaux de bord, contrats de prestations, indicateurs de performance, etc. ? L'évaluation se trouve-t-elle en mesure de réorienter, voire d'initier certaines transformations du secteur public, par exemple en établissant un lien entre les effets des politiques, l'allocation des ressources d'action aux administrations de mise en œuvre et le processus budgétaire ?

Apporter des éléments de réponse à ces questions qui traitent du statut de l'évaluation dans un contexte de modernisation de la fonction publique pré-

2. Selon notre définition, évaluer une politique publique consiste à mesurer les effets propres de cette politique et, en fonction de critères bien définis, à porter un jugement de valeur sur ces effets, qu'ils soient voulus ou non-attendus, souhaitables ou pervers, directs ou indirects, induits à court ou long terme, etc.

suppose une connaissance de l'état de la pratique évaluative dans différents pays, qu'ils se soient engagés ou non dans la NGP. Comme toute première étape de ce programme de recherche, nous nous focalisons ici sur l'analyse préalable du processus d'institutionnalisation de l'évaluation des politiques publiques. Nous cherchons ainsi à mesurer le degré d'ancrage organisationnel et procédural de l'évaluation dans dix-huit régimes démocratiques. Dans ce but, nous élaborons un indice théorique d'institutionnalisation de l'évaluation, le mesurons dans dix-huit pays retenus comme champ empirique et, finalement, explorons diverses hypothèses, dont celle d'un éventuel rôle prépondérant de la NGP lors de la diffusion de la culture évaluative, pour comprendre le fait que tous les systèmes politico-administratifs ne soutiennent pas avec la même intensité l'évaluation des politiques publiques.

### **Institutionnaliser l'évaluation des politiques publiques**

L'évaluation des politiques publiques est monnaie courante dans de nombreuses démocraties. Toutefois, en étudiant attentivement les histoires nationales et régionales de son développement, il apparaît évident que l'évaluation emprunte des cheminements divers selon les pays et, par conséquent, reflète le poids des spécificités institutionnelles et culturelles. En un mot, il s'agit d'analyser si et, le cas échéant, dans quelle mesure "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"<sup>3</sup>. Le lien postulé entre les structures, processus et acteurs d'un régime démocratique est du reste contesté par certains auteurs, dont J.-C. Thoenig qui souligne que "l'évaluation n'est pas par essence l'attribut d'un profil spécifique d'État ou de culture administrative"<sup>4</sup>, les barrières à son développement étant plus de nature pragmatique qu'institutionnelle. Pour être appropriée, l'évaluation doit être calibrée sur les problèmes concrets que rencontrent les responsables d'une action publique et apporter une plus-value à ceux-ci. D'autres auteurs au contraire soulignent l'influence prépondérante d'une culture administrative et, en particulier, de la formation en sciences sociales des fonctionnaires chargés de diffuser la pratique évaluative : "Our evidence suggests that there is an interesting degree of correspondance between the pace in the institutionalization of evaluation and the degree to which the dominant administrative philosophies, expressed in HRM (Human Resource Management's) policies, are open to innovation. Where the administrative-legal culture is clearly dominant and HRM is hesitant (or repressed) in different

3. DURAN P., "Les ambiguïtés politiques de l'évaluation", *Pouvoirs*, Paris, PUF, n° 67, 1993, p. 137.

4. THOENIG J.-C., "L'évaluation, sources de connaissances applicables aux réformes de la gestion publique", *Revue française d'administration publique*, n° 92, 1999, p. 684.

values and academic specialities, the acknowledgement of evaluation is seriously hindered”<sup>5</sup>.

La littérature présentant les expériences nationales voire sectorielles ne manque certes pas ; en revanche, les approches comparées demeurent rares<sup>6</sup>, comme en atteste la publication très récente du tout premier *International Atlas of Evaluation*<sup>7</sup>. Cet ouvrage collectif s’attache à présenter la situation de l’évaluation des politiques publiques dans vingt et un pays et trois organisations internationales (Comité d’aide au développement de l’OCDE, Union européenne et Banque mondiale) et, plus encore, à identifier les facteurs favorisant la diffusion d’une culture évaluative. Les éditeurs de l’Atlas établissent finalement un classement des pays en vertu d’un indice mesurant la maturité de la pratique évaluative qui, idéalement, se doit d’être pluraliste. Ainsi, selon J.-E. Furubo et R. Sandhal, “many of the western countries in this book have established a mature evaluation culture as an integrated element in their political and administrative systems. (...) But more important than the formal integration of evaluation in different decision-making processes, for example, the existence of laws and regulations which demand some sort of evaluative information, is the actual spread and pluralism of the evaluative culture and its openness to new ideas and impulses”<sup>8</sup>. Ci-dessous, nous confrontons cette mesure de la maturité de l’évaluation à notre propre indice théorique du degré d’institutionnalisation de l’évaluation conçu de manière parallèle au développement de l’Atlas.

### *L’institutionnalisation en débat*

Depuis la résurgence des théories institutionnalistes en sciences sociales<sup>9</sup>, le concept d’institution est à la mode et employé dans de multiples acceptions. Tel que nous le comprenons ici, dans une version très limitative d’un point de vue sociologique, le terme institution recouvre une organisation formelle ou une règle procédurale qui fournit aux acteurs un cadre assurant une certaine prévisibilité de leur comportement réciproque et, par consé-

5. BEMELMANS-VIDEC M.-L., ERIKSEN B., GOLENBERG N., “Facilitating organizational learning : human resource management and program evaluation”, in LEEUW F., RIST R., SONNICHSEN R., (eds), *Can Government Learn ? Comparative Perspective on Evaluation and Organizational Learning*, New Brunswick, Transaction Publishers, 1994, p. 179.

6. RIST R., (ed.), *Program Evaluation and the Management of Government : Patterns and Prospects across Eight Nations*, New Brunswick and London, Transaction Publishers, 1998.

7. FURUBO J.-E., RIST R., SANDAHL R., (eds), *International Atlas of Evaluation*, New Brunswick and London, Transaction Publishers, 2002.

8. FURUBO J.-E., RIST R., SANDAHL R., *op. cit.*, 2002, p. 7.

9. KOELBLE T.A., “The New Institutionalism in Political Science and Sociology”, *Comparative Politics*, 1995, vol. 27, n° 2, pp. 231-243 ; HALL P.A., TAYLOR R.C., “Political Science and the Three New Institutionalisms”, *Political Studies*, 1996, vol. 44, n° 5, pp. 936-957 ; LOWNDES V., “Varieties of New Institutionalism. A Critical Appraisal”, *Public Administration*, 1996, vol. 74, pp. 181-197.

quent, du résultat de l'action collective. L'institutionnalisation de l'évaluation correspond dès lors à une "routinisation"<sup>10</sup> du recours – attendu, si ce n'est obligé – à l'évaluation et peut être mesurée à l'aune de sa pratique effective au sein des instances politico-administratives et, plus largement, des réseaux d'action publique.

Cependant, le réflexe évaluatif au sein de l'administration peut parfois être développé sans qu'il n'existe forcément de matérialisation par des organismes politico-administratifs ou par des clauses d'évaluation explicitement ancrées dans les législations et réglementations à apprécier. Cette vision est relevée notamment par B. Dente pour lequel le succès de l'évaluation – c'est-à-dire (les effets tangibles de) son institutionnalisation – se mesure in fine par son utilisation lors de la formulation et de la mise en œuvre de politiques publiques<sup>11</sup>. Il semble donc nécessaire d'opérer une distinction entre différentes modalités d'institutionnalisation formelle et de diffusion culturelle possibles : "D'une part, les institutions classiques d'inspection et de contrôle ont été utilisées par les pouvoirs publics à des fins plus larges et plus souples ; elles se sont également elles-mêmes défendues en tentant d'adapter leurs missions initiales à des objectifs évaluatifs<sup>12</sup>. D'autre part, l'évaluation, en se répandant et en se formalisant, est devenue protéiforme. Elle s'est tantôt rigidifiée en s'institutionnalisant, tantôt au contraire assouplie jusqu'à se diluer dans un processus d'aide à la décision"<sup>13</sup>.

Sans perdre de vue cette vision disséminatrice, nous nous concentrons ici sur les éléments qui concourent à pareil essor au travers d'instances d'évaluation formelles et des sociétés nationales ou régionales qui mènent une réflexion sur la pratique et ses méthodes, destinées ensuite à essaimer au sein de l'appareil d'État. Cette approche assez restrictive nous semble pertinente car, selon l'OCDE<sup>14</sup>, entre autres, il est généralement reconnu qu'un certain degré d'institutionnalisation doit être atteint pour que l'évaluation puisse pleinement jouer son rôle dans la gestion publique. On peut en effet postuler qu'un haut degré d'institutionnalisation assure une prévisibilité de l'évaluation, au stade de la formulation de la politique publique déjà, ce qui permet de rencontrer les exigences en termes de récolte continue de données sur les actes de mise en œuvre et les effets induits. Ensuite, elle favorise l'existence d'évaluateurs qualifiés au sein de l'administration, des universités ou des cabinets de consultance. En cela elle contribue à la profession-

10. MONNIER E., *Evaluations de l'action des pouvoirs publics*, Paris, Economica, 1992, 2<sup>e</sup> édition, p. 68.

11. DENTE B., "La politique d'évaluation des politiques publiques", dans KESSLER M.-C., (dir.), *Évaluation des politiques publiques*, Paris, L'Harmattan, 1998, p. 147.

12. Sur ce point, voir POLLITT C., SUMMA H., "Evaluation and the Work of Supreme Audit Institutions : an Uneasy Relationship ?", in KESSLER M.-C., (dir.), *op. cit.*, pp. 63-82.

13. KESSLER M.-C., (dir.), *op. cit.*, p. 37.

14. OCDE, *Promouvoir l'utilisation de l'évaluation de programme*, Paris, octobre 1997.

nalisation de l'évaluation. Finalement, elle facilite la coopération entre les autorités concernées par une gouvernance à plusieurs niveaux (par exemple dans les systèmes fédéralistes) et renforce l'acceptabilité de la culture évaluative<sup>15</sup>. L'institutionnalisation de l'évaluation tend donc à éviter que des évaluations ne soient conduites que sporadiquement, empêchant l'amorce de réels processus d'apprentissage au sein des ministères et la mise en œuvre efficiente des activités d'évaluation<sup>16</sup>. De la sorte, l'institutionnalisation de l'évaluation lui confère un statut de processus social, mobilisant différents partenaires<sup>17</sup>.

### *Mesurer le degré d'institutionnalisation*

Afin de mesurer les degrés d'institutionnalisation et de pratique effective de l'évaluation, il serait évidemment opportun de recourir à des indicateurs en termes de résultats (par exemple le nombre d'évaluations, voire de méta-évaluations, réalisées)<sup>18</sup> ou de processus (par exemple le nombre d'évaluateurs actifs, le nombre d'appels d'offres publics pour des missions d'évaluation ou la part du budget public investi dans l'évaluation). Néanmoins, ces éléments statistiques ne sont pas – encore – disponibles pour permettre une analyse comparée robuste. La majorité des responsables de dispositifs nationaux d'évaluation (par exemple le Conseil national de l'évaluation en France) ne s'avèrent pas en mesure d'apporter une réponse à cette question, tant l'évaluation est polycentrique. Elle se développe dans des lieux dispersés et parfois de manière décentralisée. Finalement, ces critères perdent de leur utilité empirique dans les cas où les évaluations sont majoritairement conduites par des acteurs politico-administratifs eux-mêmes, sur un mode informel, et demeurent donc peu visibles.

Dans le même ordre d'idée, il serait tout à fait utile de recenser systématiquement les clauses d'évaluation que le législateur insère dans une loi et qui rendent obligatoire la réalisation d'une évaluation dans un délai imposé. De telles dispositions sont formulées de manière plus ou moins détaillée, allant de la simple intention d'évaluer<sup>19</sup> à la mention des conditions prati-

15. MADER L., *L'évaluation législative. Pour une analyse empirique de la législation*, Lausanne, Payot, 1985, p. 114.

16. LEEUW F., ROZENDAL P., "Policy Evaluation and the Netherland's Government : Scope, Utilization and Organizational Learning", in LEEUW F., RIST R., SONNICHSEN R., (eds), *Can Governments Learn ? Comparative Perspectives on Evaluation and Organizational Learning*, New Brunswick, Transaction Publishers, 1994, pp. 69-70.

17. PERRET B., "Institutionalise and Professionalise Evaluation : The French Case", communication présentée lors du colloque de la Société européenne d'évaluation, Rome, 29-31 octobre 1998.

18. Par exemple, en Suisse, entre 1991 et 1996, 532 titres d'évaluation ont été recensés. BUSSMANN W., *Résumé des études exécutées, mandatées ou subventionnées par les services de la Confédération pendant la période législative 1991-1995 sur l'efficience et l'effectivité de l'action étatique*, DFJP, 1996.

19. La Suisse a adopté en 1999 une disposition constitutionnelle (article 170) qui prévoit que "l'Assem-

ques de la démarche évaluative<sup>20</sup>. Tenir compte de cet indicateur imposerait toutefois de passer en revue l'ensemble de la législation et réglementation des pays comparés, ce qui semble peu réaliste : nombre de clauses ne jouissent pas d'une grande publicité et, à notre connaissance, aucun travail systématique de recensement desdites clauses évaluatives n'existe à ce jour.

Par défaut, nous nous concentrons donc sur des indicateurs situés plus en amont et indirectement liés à la pratique effective de l'évaluation certes, mais plus aisément observables et quantifiables pour un nombre élevé de pays. Nous mettons ainsi deux grandes catégories d'indicateurs (existence d'organisations formelles et d'une communauté épistémique) en évidence.

Premièrement, les structures organisationnelles particulières, comme des instances isolées ou un dispositif inter-organisationnel cohérent, revêtent évidemment une multitude de formes, mais se caractérisent toutes par une volonté marquée de représenter l'élément pivot de l'évaluation au sein de l'appareil d'État. Ces organismes peuvent être de trois types, comme l'illustre le cas de la France : parlementaire (par exemple la Mission d'évaluation et de contrôle de l'Assemblée nationale, le Comité d'évaluation du Sénat ou l'Office parlementaire d'évaluation des politiques de santé), ministériel (par exemple les dispositifs interministériels organisés autour d'un Conseil scientifique de l'évaluation puis d'un Conseil national de l'évaluation) ou administratif (par exemple la Direction de l'évaluation et de la prospective du Ministère de l'Éducation nationale). En fonction du développement de l'évaluation et du caractère unitaire ou fédéraliste de l'État, ces instances ou dispositifs peuvent se situer au niveau national (par exemple l'Office parlementaire de contrôle de l'administration en Suisse et le Contrôle fédéral des Finances) et/ou régional (par exemple la Commission externe d'évaluation des politiques publiques du canton de Genève).

Selon B. G. Peters et A. Barker, les institutions qui se développent dans une perspective d'utilisation continue s'avèrent plus efficaces que celles qui sont créées de manière ad hoc<sup>21</sup>. La régularité plutôt que l'innovation répétée encouragerait l'institutionnalisation, notamment en assurant une certaine visibilité de la pratique au sein de l'administration. Dans cette logique, il serait donc plus opportun de désigner un responsable d'évaluation avec une petite équipe permanente que de mettre en place un dispositif particulier pour chaque évaluation ponctuelle. Sur ce point, B. Perret considère a con-

---

blée fédérale veille à ce que l'efficacité des mesures prises par la Confédération fasse l'objet d'une évaluation".

20. Lorsque la Belgique dépénalise la pratique de l'interruption volontaire de grossesse ou de l'euthanasie, elle adopte dans le même temps une loi instituant une commission nationale d'évaluation de la pratique.

21. GUY PETERS B., BARKER A., *Advising West European Governments : Inquiries, Expertise and Public Policy*, Pittsburgh, University Press, 1993, p. 12.

trario que les instances ad hoc, souvent critiquées pour leur lourdeur, favorisent une socialisation et une diffusion des concepts. Il interprète ces organisations comme des lieux d'apprentissage collectifs dont les membres deviennent ensuite les meilleurs "propagandistes" des résultats<sup>22</sup>.

Au-delà de l'existence ou non d'une instance formelle, il s'agirait aussi de mesurer les missions et les prérogatives de ces instances ainsi que leurs ressources disponibles et, finalement, leur centralité dans le dispositif institutionnel global. Ce travail présuppose la réalisation d'études de cas qualitatives qui présentent ces éléments en détail dans toute leur singularité<sup>23</sup>.

Deuxièmement, l'existence d'une communauté épistémique (nationale)<sup>24</sup> représente, à notre avis, un élément définitionnel important de l'institutionnalisation de l'évaluation. J.-E. Furubo, souligne que, depuis les années 1990, de nouveaux moyens d'information et de communication et d'organisations internationales (dont par exemple la Société européenne d'évaluation) catalysent la constitution d'un véritable réseau professionnel. Les élites intellectuelles et politico-administratives de différents pays entretiennent ainsi plus de relations au sein d'arènes informelles voire virtuelles<sup>25</sup>. À cet égard, les concepts d'arène et de forum développés par B. Jobert s'avèrent particulièrement utiles pour identifier les lieux et chemins empruntés par l'institutionnalisation de l'évaluation. Selon cet auteur, les forums sont des lieux de discussion et de débat sur le sens des politiques publiques tandis que les arènes constituent des espaces de négociation des compromis institutionnalisés<sup>26</sup>.

Au niveau national, les principales manifestations de l'existence d'une communauté épistémique sont les clubs<sup>27</sup>, les réseaux<sup>28</sup> et les sociétés nationales d'évaluation<sup>29</sup> qui regroupent les évaluateurs et commanditaires institutionnels du territoire donné. Les deux premières étant le plus souvent

22. PERRET B., *L'évaluation des politiques publiques*, Paris, La Découverte, 2001, p. 16.

23. Pour une telle analyse des dispositifs et processus d'institutionnalisation de l'évaluation des politiques en Belgique, en France, en Suisse et aux Pays-Bas, voir JACOB S., *L'institutionnalisation de l'évaluation des politiques publiques en Europe. Etude comparée des dispositifs institutionnels en Belgique, en France, en Suisse et aux Pays-Bas*, Thèse de doctorat inédite, Bruxelles, Université libre de Bruxelles, 2004.

24. HAAS P., "Introduction : Epistemic Communities and International Policy Coordination", *International Organization*, 1992, vol. 46, n° 1, pp. 1-35.

25. FURUBO J.-E., RIST R., SANDAHL R., *op. cit.*, 2002, p. 18.

26. JOBERT B., "Rhétorique politique, controverses scientifiques et construction des normes institutionnelles : esquisse d'un parcours de recherche", dans FAURE A., POLLET G., WARIN P., (dir.), *La construction du sens dans les politiques publiques. Débats autour de la notion de référentiel*, Paris, L'Harmattan, 1995, pp. 19-23.

27. Club Cambon ou Club de l'évaluation en France par exemple.

28. "Netzwerk suisse" par exemple.

29. Il existe de nombreuses sociétés nationales d'évaluation. À titre d'exemple mentionnons la Société canadienne d'évaluation ([www.evaluationcanada.ca](http://www.evaluationcanada.ca)), la Société italienne d'évaluation ([www.valutazioneitaliana.it](http://www.valutazioneitaliana.it)) et la Société française d'évaluation ([www.sfe.asso.fr](http://www.sfe.asso.fr)).



informelles et parfois limitées dans le temps, nous nous concentrons essentiellement sur les sociétés nationales comme indicateur mesurable. Au-delà de leurs activités traditionnelles (par exemple les colloques, la formation continuée ou le carnet d'adresses), certaines sociétés d'évaluation rédigent des standards de qualité, publient un bulletin d'information destiné à leurs membres et/ou éditent une revue scientifique. Il apparaît donc que ces sociétés sont plus ou moins actives dans leur mission de promotion et de développement de l'évaluation des politiques publiques.

Les revues scientifiques spécialisées en évaluation (par exemple *Evaluation and Program Planning*, *Evaluation Review*, *Evaluation Practice*, *Evaluation : The International Journal of Theory, Research and Practice*) constituent, pour la communauté des évaluateurs, un forum permettant la diffusion d'expériences nationales et/ou internationales, l'élaboration d'un corpus théorique et méthodologique commun, etc. En fait, la communauté des évaluateurs regroupe une variété de disciplines (science économique, science politique, psychologie, etc.), de statuts professionnels (universitaires, consultants privés et fonctionnaires) et d'expériences sectorielles (santé publique, énergie, emploi, etc.). Les revues essaient donc de refléter cette diversité et de la dépasser en confrontant différentes pratiques, théories ou méthodes évaluatives. Afin d'esquisser un indice d'institutionnalisation de l'évaluation qui rend compte de ces réflexions scientifiques, il serait intéressant de réaliser une enquête bibliométrique de ces revues et de la participation aux colloques internationaux (par exemple en fonction de la nationalité des auteurs d'articles, des expériences évaluatives présentées, etc.). De toute évidence, pareille entreprise dépasse le cadre de cet article introductif.

Les standards professionnels déterminent quant à eux des lignes de conduite garantissant le respect de normes de qualité en termes de déontologie, de faisabilité pratique, d'utilité politique et de rigueur scientifique des évaluations réalisées au concret<sup>30</sup>. Leur utilisation témoigne donc d'une volonté de réaliser des évaluations de qualité<sup>31</sup>. Certains proposent même de recourir à ces standards pour procéder à une certification des évaluateurs et évaluatrices. Actuellement, une des pratiques courantes d'une société d'évaluation consiste à rédiger ses propres standards. De cette manière, elle peut adapter ces règles à certains particularismes nationaux.

En mobilisant uniquement les éléments tangibles de l'institutionnalisation de l'évaluation des politiques publiques, même si nous sommes bien

---

30. Par exemple les *Joint Committee on Standards for Educational Evaluation* américains. Pour une présentation comparée des standards, voir NADEAU M.-A., *L'évaluation de programme. Théorie et pratique*, Presses de l'Université Laval, 1988, pp. 141-165. WIDMER T., LANDERT C., BACHMANN N., *Standards d'évaluation recommandés par la Société suisse d'évaluation (SEVAL)*, 1999. Disponible en ligne : [www.seval.ch](http://www.seval.ch).

31. NEWMAN D., BROWN R., "Violations of Evaluation Standards : Frequency and Seriousness of Occurrence", *Evaluation review*, 1992, vol. 16, n° 3, p. 220.

conscients de la simplification héroïque que cache cette mesure, nous construisons un indice théorique<sup>32</sup> dont le tableau 1 ci-dessous présente les différents éléments, et les compare aux dimensions constitutives de l'indice de maturité proposé par J.-E. Furubo et ses collègues dans l'*International Atlas of Evaluation* mentionné plus haut.

Ces deux indices permettent une comparaison internationale des degrés d'institutionnalisation et de maturité de la pratique évaluative en se basant sur une analyse empirique de leurs composantes respectives (auxquelles des points sont attribués).

### ***Positionnement et typologie de dix-huit pays***

À titre d'exemple<sup>33</sup>, le degré d'institutionnalisation de la France est égal à 6 (le maximum théorique étant 9) car elle dispose d'une instance exécutive (le Conseil national de l'évaluation, 2 points) et d'une instance législative (la Mission d'évaluation et de contrôle de l'Assemblée nationale, 2 points). Par contre, la Cour des comptes française ne dispose pas de prérogatives particulières en matière d'évaluation (0 point). Au niveau de la communauté épistémique, il existe une Société française d'évaluation (1 point) qui a mandaté un groupe de travail chargé de réfléchir à la question des standards de qualité (1 point) mais qui ne dispose pas de revue (0 point). En termes de maturité de la pratique, la France obtient un score de 11 (le maximum théorique étant 18) selon l'appréciation faite par C. Fontaine et E. Monnier dans l'*Atlas international*<sup>34</sup>. Ce classement tient compte des critères suivants : l'évaluation a lieu dans des domaines variés (2 points) avec le soutien d'évaluateurs internes (1 point) dans un cadre pluraliste (1 point). Il existe un discours sur l'évaluation (1 point) et des organisations professionnelles (2 points). Finalement, l'évaluation est institutionnalisée au sein de l'exécutif (2 points) et auprès du Parlement (1 point), de même qu'elle est pratiquée par la Cour des comptes (1 point).

La figure 1 ci-dessous présente les mesures empiriques des deux indices théoriques pour l'ensemble des dix-huit pays communs aux deux études menées en parallèle<sup>35</sup>.

32. JACOB S., VARONE F., *Évaluer l'action publique : état des lieux et perspectives en Belgique*, Gand, Academia Press, 2003, pp. 115 ss.

33. Entre parenthèses, nous indiquons les points accordés au pays analysé en fonction de l'existence ou non des différents éléments constitutifs des indices.

34. FONTAINE C., MONNIER E., "Evaluation in France", in FURUBO J.-E., RIST R., SANDAHL R., *op. cit.*, 2002, pp. 63-75.

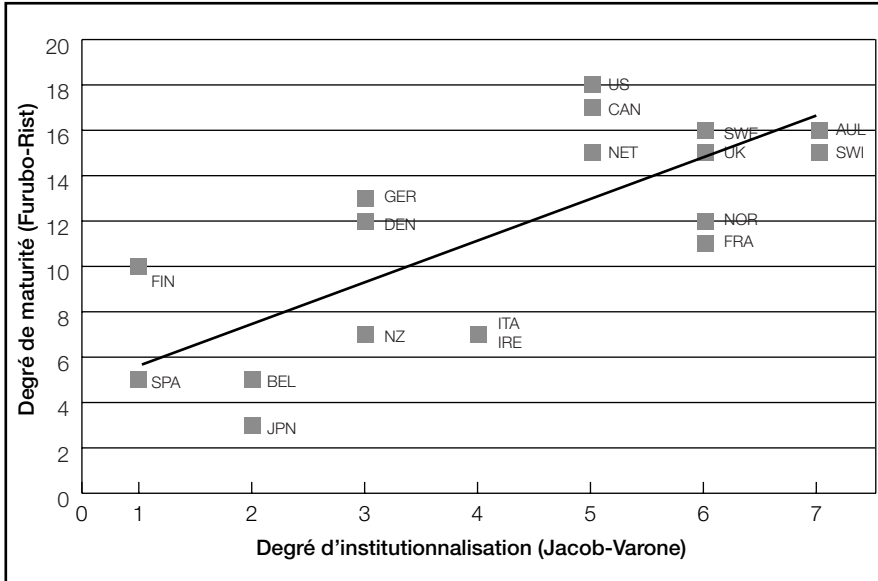
35. À l'exception de la Belgique (que nous avons codée nous-mêmes), les pays qui ne sont étudiés que par une des deux études ne sont pas repris ici. Il s'agit de la Grèce, de l'Islande, du Luxembourg, du Portugal (traités par Jacob-Varone mais pas par Rist-Furubo) et de la Chine, Zimbabwe et Corée du Sud (traités par Rist-Furubo mais pas par Jacob-Varone). Le chapitre de l'Atlas consacré à la Suisse est

Tableau 1 : Composantes des indices théoriques des degrés de maturité et d'institutionnalisation de l'évaluation

	Institutionnalisation selon Jacob-Varone (indice de 0 à 9)	Maturité de la pratique selon Rist et Furubo (indice de 0 à 18)
<b>Instances d'évaluation</b>		L'évaluation est répandue dans de nombreux secteurs de politiques publiques (de 0 à 2)
	Instances d'évaluation auprès de l'Exécutif (de 0 à 2)	Degré d'institutionnalisation – Gouvernement (de 0 à 2)
	Instances d'évaluation auprès du Parlement (de 0 à 2)	Degré d'institutionnalisation – Parlement (de 0 à 2)
	Pratique de l'évaluation au sein de la Cour des comptes (de 0 à 2)	Pratique de l'évaluation au sein de la Cour des comptes (de 0 à 2)
		Soutien d'évaluateurs (nationaux) pluridisciplinaires (de 0 à 2)
<b>Règles Forum Pratiques</b>	Société nationale d'évaluation (de 0 à 1)*	Organisation professionnelle (de 0 à 2)
	Revue scientifique (de 0 à 1)	Discours national concernant l'évaluation (de 0 à 2)
	Standards de qualité (de 0 à 1)	Pluralisme des évaluations ou des évaluateurs au sein d'un secteur de politiques publiques (de 0 à 2)
		Prise en compte par l'évaluation des effets et de la mise en œuvre des politiques publiques (de 0 à 2)

\* Afin de pouvoir classer les différents pays considérés, nous attribuons donc plus d'importance aux facteurs formels qu'informels.

**Figure 1 : Mesure empirique des degrés de maturité et d'institutionnalisation de l'évaluation dans dix-huit pays**



De toute évidence, les deux indices sont corrélés entre eux<sup>36</sup>. De plus, et comme le montre le tableau 2 ci-dessous, trois catégories de pays se distinguent aisément. À un extrême, nous trouvons les pays qui ont un degré d'institutionnalisation élevé et qui, conjointement, présentent un degré de maturité important (pays pionniers). À l'autre extrême se situent les pays qui affichent une pratique évaluative peu développée de même qu'un faible degré d'institutionnalisation (pays retardataires). Le groupe de pays intermédiaires se caractérise quant à lui par une certaine maturité de la pratique, bien que celle-ci ne repose pas formellement sur une institutionnalisation élevée. Soulignons encore qu'aucun pays – parmi ceux analysés ici – ne s'est engagé dans un processus d'institutionnalisation poussée sans pour autant que cela ne se traduise ensuite, ou de manière concomitante, par une croissance de la maturité de l'évaluation. De ce point de vue, il semble donc que les

contesté par plusieurs professionnels helvétiques de l'évaluation (dont par exemple : JANETT D., "Evaluation wird als Kontrollinstrument des Bundesparlaments immer wichtiger", *Leges : législation et évaluation*, n° 1, 2003, pp. 165-167). En conséquence, nous avons recodé le degré de maturité de la Suisse sur la base de nos connaissances antérieures et de notre étude de cas détaillée (cf. JACOB S., VARONE F., *op. cit.*, 2003).

36. Il existe une corrélation entre ces deux indices (Pearson Correlation = 0,724\*\* – Sig. (2-tailed) =,001). La moitié de la variance est expliquée (Rsq = 0,5241). La courbe de tendance qui cherche à expliquer la maturité ( $y = 1,7304$ ) est plus grande que celle qui cherche à expliquer l'institutionnalisation ( $y = 0,3029$ ). Au stade préalable, où nous nous trouvons, nous pouvons en déduire que l'institutionnalisation influence la maturité. Toutefois, cette affirmation doit être nuancée en raison de la différence d'échelle entre les deux indices (de 0 à 9 d'une part et de 0 à 18 d'autre part) et d'une relative autocorrélation entre les éléments constitutifs des deux indices.

instances et règles formelles d'institutionnalisation de l'évaluation ne demeurent pas des coquilles vides ; au contraire, elles déploient des effets sur la pratique effective de l'évaluation des politiques publiques.

Une des interrogations qui ne manquent pas de surgir à la lecture de la figure 1 concerne l'influence prépondérante d'une dimension sur l'autre : est-ce l'institutionnalisation qui influence la maturité (une fois que les institutions formelles et règles institutionnelles existent, une plus grande demande d'évaluations émerge, ce qui conduit à augmenter la maturité) ou bien la maturité qui influence l'institutionnalisation (le grand nombre d'évaluations réalisées engendre la création d'institutions pour en encadrer et formaliser la pratique) ? À ce stade, il est difficile de déterminer la prépondérance ou l'influence directe de l'un de ces facteurs. Au regard de la typologie présentée dans le tableau 2, il apparaît néanmoins que l'expérimentation de l'évaluation au quotidien et, par ce biais, l'émergence graduelle d'une culture évaluative représentent une condition nécessaire, mais pas suffisante, pour un processus d'institutionnalisation de cet outil.

### **Conditions d'une institutionnalisation de l'évaluation**

Les études de cas nationales et sectorielles disponibles à ce jour identifient plusieurs facteurs influençant le degré d'institutionnalisation de l'évaluation. Vu la nature exploratoire de notre analyse comparée, nous reprenons à notre compte ces hypothèses, souvent formulées pour un pays et une période bien définis, et les appliquons systématiquement aux dix-huit démocraties de notre champ empirique. Outre les effets postulés de la NGP sur la pratique évaluative, nous retenons cinq autres arguments qui soulignent le rôle potentiel de variables contextuelles (situation fiscale et pression externe) et institutionnelles (régime majoritaire, instances de contrôle et structure fédéraliste) sur le niveau de développement de l'évaluation des politiques publiques.

### ***Nouvelle Gestion Publique : un frein à l'évaluation ?***

En premier lieu, la question posée est de savoir si, et dans quelle mesure, la NGP représente un facteur facilitant ou inhibant l'évaluation. Cette interrogation s'avère d'autant plus pointue que les programmes de modernisation inspirés par les principes de la NGP ne reprennent généralement pas, en tout cas pas comme une priorité d'action explicite, l'évaluation des politiques publiques<sup>37</sup>. Au contraire, les tenants de l'évaluation ont souligné à de nom-

---

37. Sur ce point, l'exemple récent du Japon qui a adopté, en 2001, une loi sur l'évaluation des politiques publiques semble constituer une exception. Cependant, la mise en œuvre de celle-ci est trop récente pour

Tableau 2 : Répartition des dix-huit pays en trois catégories

		<b>DEGRÉ D'INSTITUTIONNALISATION</b>	
<b>DEGRÉ DE MATURITÉ</b>	Élevé (si l'indice est supérieur ou égal à 11)	Élevé (si l'indice est supérieur à 4) Angleterre, Australie, Canada, États-Unis, France, Norvège, Pays-Bas, Suède, Suisse	Bas (si l'indice est inférieur ou égal à 4) Allemagne, Danemark
	Bas (si l'indice est inférieur à 11)		Belgique, Espagne, Irlande, Japon, Nouvelle-Zélande, Italie, Finlande

breuses reprises que la NGP tend à se focaliser plutôt sur l'efficacité interne de l'administration, en rationalisant (les coûts de) la production des prestations administratives ("outputs") par un service administratif particulier, que sur la mesure et le jugement des effets propres ("outcomes", impacts) des politiques conduites, avec une coordination plus ou moins proactive, par plusieurs unités administratives, organismes parapublics voire associations privées.

D'un point de vue purement chronologique, il apparaît évident que les pionniers de l'évaluation se sont attachés à développer cette pratique avant même que le courant néo-libéral de la NGP ne propose ses solutions. Ainsi en est-il de la situation aux États-Unis, au Canada, au Royaume-Uni, en Suède et en Allemagne, soit les pays initiateurs des premières vagues d'institutionnalisation de l'évaluation<sup>38</sup>. Pour ces systèmes politico-administratifs, l'enjeu est donc de savoir si l'introduction postérieure de la NGP a renforcé ou au contraire limité la pratique évaluative. Aux États-Unis par exemple, l'introduction en 1993 du *Government Performance and Results Acts (GPRA)* s'est traduite *de facto* par un déclin significatif des ressources allouées et (des compétences) du personnel affecté à l'évaluation selon l'analyse quantitative des quatorze agences américaines chargées de missions d'évaluation et comptabilisant 98 % du budget fédéral 1994 consacré à l'évaluation<sup>39</sup>.

Pour les États qui n'avaient pas pérennisé l'évaluation avant de se lancer dans d'éventuelle réforme inspirée de la NGP, l'expérience semble largement similaire. Le cas emblématique de la Nouvelle-Zélande illustre à dessein le fait que les nouveaux gestionnaires publics s'engagent à atteindre des critères de performance portant sur les produits administratifs fournis par leur service respectif. Par contre, il ne souhaitent nullement être tenus responsables des effets induits par lesdites prestations administratives auprès des destinataires finaux des politiques car, et c'est une évidence, ces effets dépendent non seulement des outputs réalisés par leur service, mais aussi de la qualité de la programmation normative de la politique en soi, de sa mise en œuvre conjointe par d'autres unités administratives et de variables contextuelles, facteurs sur lesquels ils n'ont aucune prise directe et dont, très logiquement, ils ne sauraient être tenus responsables en cas d'effets déficitaires<sup>40</sup>. En un mot, tant la finalité étroitement managériale de la NGP (amé-

---

pouvoir en apprécier les effets concrets. KUDO H., "Entre modèle de 'gouvernance' et loi sur l'évaluation des politiques : le nouveau management public au Japon", *Revue internationale des sciences administratives*, 2003, vol. 70, n° 4, pp. 567-592.

38. Sur les vagues d'institutionnalisation, voir DERLIEN H.-U., "Genesis and Structure of Evaluation Efforts", in RIST R., *op. cit.*, 1990, pp. 147-175.

39. WARGO M., "The Impact of Federal Government Reinvention on Federal Evaluation Activity", *Evaluation Practice*, 1995, vol. 16, n° 3, pp. 227-237.

40. Cf. POLLITT C., BOUCKAERT G., *Public Management Reform. A Comparative Analysis*, Oxford. Oxford University Press, 2000, p. 88 et pp. 110ss.

liorer l'efficacité interne) que le jeu stratégique des gestionnaires administratifs (contrôler ses propres performances) contribuent plutôt à diminuer le poids de l'évaluation.

Les résultats d'autres études nationales laissent peu d'ambiguïté sur le fait que la NGP tend, en règle générale, à réduire les efforts d'évaluation des effets des politiques publiques. H-U. Derlien et R. Rist résument cet enseignement de la sorte : "Impacts indicators are sacrificed for outputs measures as evident in those countries epousing New Public Management. Performance indicators are often at best outputs measures, leaving out considerations of distributive effects"<sup>41</sup>. L'hypothèse de recherche qui découle logiquement de ces constats et qui mérite d'être testée sur un large ensemble de pays stipule donc, dans une formulation volontairement simplifiée, que la NGP a une influence plutôt négative sur le degré d'institutionnalisation et, plus encore, sur la maturité de l'évaluation des effets des politiques publiques.

### *Déficits budgétaires : évaluer pour mieux couper ?*

Deuxièmement, l'évaluation peut être instrumentalisée pour opérer et légitimer des réductions budgétaires dans certains secteurs, en fonction d'analyses coûts-bénéfices des politiques publiques concernées. En situation de récession économique et de déficits publics récurrents, les informations livrées par l'évaluation se voient parfois valorisées lors du processus budgétaire ou, à tout le moins, lors de l'engagement de diverses dépenses publiques. Cette stratégie de rationalisation financière, pilotée dans les années 1980 par les ministères des finances et d'autres gardiens du budget, caractérise notamment les situations de l'Australie, du Canada, du Royaume-Uni, de la Norvège et des Pays-Bas<sup>42</sup>.

Le cas de la procédure de reconsidération initiée aux Pays-Bas conjointement par le Ministère des Finances et la Cour des comptes atteste de la plausibilité de cette hypothèse. Durant les années 1970, la plupart des économistes hollandais estiment en effet que seul un déficit public évoluant autour de 4 % ou 5 % est acceptable<sup>43</sup>. C'est pour cette raison que lorsque le déficit s'accroît de manière considérable au début des années 1980, le Gouvernement décide de réagir afin de juguler le déficit qui atteint des niveaux "intolérables". Parmi les solutions préconisées et mises en place pour économiser des fonds publics figure la procédure de reconsidération, considé-

41. DERLIEN H.-U., RIST R., "Conclusion. Policy Evaluation in International Comparison", in FURUBO J.-E., RIST R., SANDAHL R., *op. cit.*, 2000, p. 450.

42. DERLIEN H.-U., RIST R., *op. cit.*, 2002, p. 443.

43. ANDEWEG R., IRWIN G., *Dutch Government and Politics*, London, Macmillan, 1993, p. 209.



rée par beaucoup comme les prémices de l'évaluation aux Pays-Bas<sup>44</sup>. Dès lors, l'hypothèse générale à tester stipule que plus le déficit budgétaire est important, plus le degré d'institutionnalisation de l'évaluation est élevé.

### ***Contraintes extérieures : de l'obligation réglementaire à l'initiative propre ?***

L'institutionnalisation de l'évaluation peut être sous-tendue par une pression externe au pays concerné. En particulier, des exigences de nature réglementaire lui sont imposées pour qu'il mène une démarche évaluative, comme cela est le cas pour les États-membres de l'Union européenne bénéficiant des Fonds structurels ou pour les pays récipiendaires d'une aide financière octroyée, de manière conditionnelle, par certains organismes internationaux dont la Banque mondiale par exemple. Rendre formellement des comptes à son financeur (selon les fameux principes de "accountability" et "value-for-money"), en recourant à une évaluation des effets induits par les politiques de développement subventionnées, correspond dès lors à une première étape dans la pratique évaluative et le développement d'un réflexe évaluatif.

À titre illustratif, notons que près de 400 évaluations ont ainsi été réalisées de 1996 à 1998 dans le cadre des Fonds structurels. Ce qui témoigne d'une diffusion spectaculaire de cet outil, surtout dans les pays du Sud de l'Europe. Selon M. Pazos et E. Zapico-Goni, le cas de l'Espagne est emblématique à cet égard : "With Spain being one of the greatest recipients of the European funds, and the evaluation of these funds being the most important factor for launching evaluation practices and infrastructure, the European Commission has had a significant influence on the development of evaluation in Spain"<sup>45</sup>. Au-delà de l'impulsion initiale donnée par l'Union européenne, plusieurs pays s'approprient ensuite l'évaluation pour, d'initiative propre cette fois, l'appliquer systématiquement à leurs politiques internes, dans une visée plus managériale et politique que réglementaire<sup>46</sup>. L'hypothèse de recherche qui en découle soutient donc que plus la pression externe à procéder à des évaluations est importante, plus le degré d'institutionnalisation est élevé.

---

44. BEMELMANS-VIDEC M.-L., "Dutch Experience in the Utilization of Evaluation Research : the Procedure of Reconsideration", *Knowledge in Society*, 1989, n° 2, pp. 31-49.

45. PAZOS M., ZAPICO-GONI E., "Program Evaluation in Spain : Taking Off at the Edge of the Twenty-First Century ?", in FURUBO J.-E., RIST R., SANDAHL R., *op. cit.*, 2002, p. 297.

46. VANDENDORPE L., "De l'évaluation des politiques régionales à une politique régionale de l'évaluation en Wallonie", dans DE VISSCHER C., VARONE F., (eds), *Evaluer les politiques publiques. Regards croisés sur la Belgique*, Louvain-la-Neuve, Academia Bruylant, 2001, pp. 67-78.

### ***Régime démocratique majoritaire : l'évaluation comme levier de l'opposition ?***

Certains auteurs avancent que l'évaluation représente un levier d'action que peut activer l'opposition parlementaire à l'encontre des politiques gouvernementales. Interprétée comme un instrument de contrôle, si ce n'est carrément de contre-pouvoir, l'évaluation aurait ainsi tendance à être plus institutionnalisée dans les démocraties majoritaires (soit le modèle de Westminster ou un régime présidentiel) que dans les démocraties dites de consensus, de négociation ou de concordance<sup>47</sup>. Ces dernières connaissent en effet des coalitions gouvernementales qui ne permettent pas d'imputer clairement le succès ou l'échec d'une politique, tel qu'objectivé par une évaluation, à un parti politique en particulier.

De plus, le respect d'accords gouvernementaux ou programmes de législature (en Belgique par exemple) et le principe de collégialité qui encadre le fonctionnement de certains exécutifs (en Suisse par exemple) limitent certainement le jeu stratégique des partis vis-à-vis de l'évaluation et de son institutionnalisation. Nombre d'analystes soulignent par exemple que la participatie et le système de coalition gouvernementale belges représentent un frein à la pratique évaluative<sup>48</sup>. Selon N. Stame, la démocratie consociative ("consociativismo") représente aussi un frein à l'institutionnalisation de l'évaluation en Italie<sup>49</sup>. Ainsi, nous postulons que plus le régime politique est majoritaire, plus le degré d'institutionnalisation est élevé. Nous pouvons préciser plus avant cette hypothèse, en soulignant que, sur base de l'argument de contre-pouvoir présenté ci-dessus, le degré d'institutionnalisation est théoriquement plus élevé dans les régimes présidentiels (séparation des pouvoirs) que dans les systèmes à la Westminster (fusion des pouvoirs).

### ***Indépendance des organes de contrôle : évaluation comme nouvel outil ?***

Cinquièmement, et toujours dans le même ordre d'idée, le degré d'indépendance des organes de contrôle traditionnels (soit la Cour des comptes ou des équivalents fonctionnels) vis-à-vis du pouvoir exécutif est retenu comme facteur facilitant l'institutionnalisation de l'évaluation. Si ces instances jouissent d'une autonomie institutionnelle élevée et des ressources nécessaires, y compris en termes d'expertise méthodologique, alors elles chercheront à s'appropriier l'outil évaluatif, ne serait-ce qu'à des fins de renforcement et

47. LIJPHART A., *Patterns of Democracy : Government Forms and Performance in Thirty-six Countries*, New Haven, Yale University Press, 1999.

48. DESTEXHE A., ERALY A., GILLET E., *Démocratie ou participatie ?*, Bruxelles, Labor, 2003 ; ERALY A., *Le pouvoir enchaîné. Être ministre en Belgique*, Bruxelles, Labor, 2002.

49. STAME, N., "Evaluation in Italy : An Inverted Sequence from Performance Management to Program Evaluation ?", in FURUBO J.-E., RIST R., SANDAHL R., *op. cit.*, 2002, p. 274.

(re)légitimation de leur propre organisation, et à se faire reconnaître comme un acteur du dispositif institutionnel d'évaluation des politiques publiques. Malgré une volonté affichée d'apparaître concernées par les question évaluatives, le degré d'implication des Cours des comptes est toutefois très différent selon les pays.

À l'instar d'autres organisations très anciennes, les Cours des comptes<sup>50</sup> voient leurs rôles et missions évoluer. Certaines s'engagent clairement dans une démarche où l'évaluation est au cœur de leur travail (par exemple aux Pays-Bas et en Suisse) tandis que d'autres, soucieuses de préserver l'image de rigueur et de respectabilité que leur procurent leurs missions traditionnelles de contrôle (régularité, conformité, légalité) des deniers publics, tentent d'accommoder leur travail avec quelques aspects empruntés à l'évaluation des politiques publiques (par exemple en France et en Belgique). De la sorte, il semble que si développement de la pratique évaluative il y a, alors les acteurs parlementaires mobiliseront ces organes de contrôle pour réaliser des évaluations de l'action gouvernementale. À ce sujet, H.-U. Derlien et R. Rist formulent l'hypothèse générale suivante<sup>51</sup> : "The more the central auditing unit is independent from the executive branch, and even attached to Parliament, the more parliaments may be tempted to use these institutions as a power basis and as a source of counter-advice vis-à-vis even a party government."

### *Structure fédéraliste : le laboratoire de l'évaluation ?*

La dernière hypothèse que nous retenons ici soutient que si le système politique est fédéraliste, alors le degré d'institutionnalisation de l'évaluation est élevé. D'une part, l'évaluation rend effectif le suivi de la mise en œuvre de lois fédérales par les entités fédérées, comme le veut la séparation des tâches caractérisant les fédérations dites fonctionnelles (la Suisse ou l'Allemagne par exemple). D'autre part, le renforcement de l'évaluation pourrait aussi résulter d'un processus d'émulation et d'apprentissage entre les entités fédérées, certaines d'entre-elles assumant un rôle d'expérimentateurs de l'évaluation au sein du "laboratoire fédéraliste".

Un cas parlant à cet égard est celui de la Belgique où l'évaluation connaît des développements extrêmement nuancés entre les différents niveaux de pouvoir (fédéral, régional ou communautaire). Dans cette fédération juridictionnelle, qui repose sur une répartition des compétences de politiques publiques entre entités publiques, la Wallonie est assurément pionnière en

---

50. La Suisse ne dispose pas, à proprement parler, d'une Cour des comptes. C'est sur son équivalent fonctionnel, le Contrôle fédéral des Finances, que nous nous sommes attardés dans cette partie.

51. DERLIEN H.-U., RIST R., *op. cit.*, p. 445.

matière d'institutionnalisation de la pratique évaluative. En effet, alors que l'évaluation connaît quelques balbutiements en Flandre, en Région de Bruxelles-Capitale, en Communauté française ou au niveau fédéral, le Gouvernement wallon a érigé l'évaluation comme une des priorités de son programme d'action formalisé dans un texte intitulé *Le Contrat d'avenir pour la Wallonie*. En terme institutionnel, cet engagement se concrétise à travers la constitution d'une Société wallonne de l'évaluation et de la prospective (SWEP) et de la création d'un Institut wallon de l'évaluation, de la prospective et de la statistique (IWEPS). Les autres niveaux de pouvoir ne demeurent pas insensibles aux expériences entreprises en Wallonie et des processus d'isomorphisme institutionnel pourraient s'observer à moyen terme.

### **Résultats empiriques**

Afin de tester nos six hypothèses de travail, nous avons procédé à des tests statistiques (coefficients de corrélation) ainsi qu'à une analyse quali-quantitative comparative (AQQC basée sur l'algèbre booléenne<sup>52</sup>). Les résultats de ces deux démarches nous poussent à conclure que les liens postulés ci-dessus, et dont la plausibilité pour certaines histoires nationales de l'évaluation semble pourtant évidente, ne résistent pas à un test systématique sur une base empirique large. En effet, aucun des coefficients calculés ne s'avère significatif, de même que les équations minimales résultant de l'analyse AQQC ne nous permettent pas de répartir de manière parcimonieuse, dans des catégories regroupant plusieurs pays caractérisés par une constellation de conditions identiques, les dix-huit démocraties retenues dans notre échantillon. En un mot, notre étude exploratoire tend à rejeter l'idée selon laquelle un nombre limité de trajectoires communes, voire de types idéaux, permettrait de cerner les processus d'institutionnalisation et, par-là, d'expliquer les différents niveaux de pratique de l'évaluation dans les pays occidentaux. Dit autrement, les liens postulés par nombre d'auteurs entre structures politico-administratives, variables contextuelles et NGP demeurent, pour le moins, ténus et contingents.

Cet enseignement empirique dépend naturellement de la manière dont nous avons conçu et opérationnalisé les variables indépendantes (institutionnelles et contextuelles)<sup>53</sup> et dépendantes (degrés de maturité et d'insti-

52. Voir le numéro spécial de la RIPC consacré à l'analyse quali-quantitative comparative : *Revue Internationale de Politique Comparée*, 2004, vol. 11, n° 1.

53. Pour opérationnaliser nos variables indépendantes, sur la période 1970-2000, nous avons recouru à des données et indices traditionnels de la littérature : les données de l'OCDE sur le développement de la NGP et la prise en compte d'informations sur les performances non-financières dans les budgets (hypothèse 1), les déficits budgétaires en pourcentage du Produit National Brut (hypothèse 2), le ratio entre la population et les montants alloués au titre de fonds régionaux par l'Union européenne (hypothèse 3), les indices définis par A. Lijphart ("executive parties dimension", "executive dominance index") sur les

tutionnalisation) de notre cadre théorique. Afin de mieux calibrer notre modèle d'analyse, des modifications devraient être apportées à différents niveaux. Par exemple, il s'agirait de retenir non pas un état des lieux par pays à un moment donné mais de distinguer, au sein d'un même pays, différents dispositifs institutionnels selon différentes périodes. Ceci présupposerait une étude historique plus circonstanciée des dispositifs qui se suivent. En France par exemple, nous observons que le dispositif "Rocard" structuré autour du Conseil scientifique de l'évaluation (1990) est remplacé par un dispositif centré autour du Conseil national de l'évaluation (1998) lorsque Lionel Jospin est Premier ministre. Des dispositifs interministériels aux finalités et à l'architecture différentes constitueraient ainsi les phénomènes politiques à décrire et à expliquer. Il serait dès lors opportun d'analyser plus en détail, non pas une situation statique, mais bien plus le processus dynamique d'émergence et de transformation des institutions qui encadrent la pratique évaluative. Un tel programme de recherche renvoie aux analyses comparatives historiques qui se penchent sur les points de bifurcation lors d'innovations institutionnelles, mais aussi sur la sédimentation, la conversion et la recombinaison institutionnelles<sup>54</sup>.

De plus, il serait pertinent d'ajouter de nouvelles variables explicatives, plus directement orientées vers les acteurs de l'évaluation et de son institutionnalisation. Ainsi, en est-il des partis politiques qui peuvent soutenir ou s'opposer à un renforcement de l'évaluation puisque, dans le cas français, les dispositifs érigés lorsque la gauche est au pouvoir sont délaissés dès que la droite gouverne. Pareille situation est actuellement observable avec le deuxième dispositif interministériel structuré autour du Conseil national de l'évaluation dont les membres n'ont plus été remplacés depuis l'arrivée de J-P. Raffarin à Matignon. Cette opposition entre droite et gauche est constatée non seulement en France<sup>55</sup> mais dans d'autres pays. Soulignons toutefois que des partis tant de gauche (progressistes) que de droite (conservateurs) ont de facto soutenu, avec des stratégies différentes certes, le développement de l'évaluation<sup>56</sup>.

---

démocraties majoritaires et consensuelles (hypothèse 4), la mesure de l'indépendance des Cours de comptes par l'INTOSAI (hypothèse 5) et, finalement, l'indice proposé par A. Lijphart ("federal-unity dimension") pour qualifier le degré de fédéralisme (hypothèse 6), LIJPHART A., *op. cit.*, 1999

54. Cf. THELEN C., "How Institutions Evolve : Insights from Comparative Historical Analysis", in MAHONEY J., RUESCHENMEYER D., (eds), *Comparative Historical Analysis in the Social Sciences*, New York, Cambridge University Press, 2003, pp. 208-240. Pour une excellente synthèse de ce courant, voir BOYER R., "Les analyses historiques comparatives du changement institutionnel : quels enseignements pour la théorie de la régulation ?", dans *L'Année de la régulation 2003 : Economie, Institutions, Pouvoirs*, 2003, vol. 7, pp. 167-203.

55. L. Rouban formule l'hypothèse suivante pour la France : "La sensibilité de gauche serait favorable au développement d'une évaluation qui pourrait remplir des fonctions d'explicitation et de pilotage des politiques publiques, les représentations de droite favorisant une vision plus traditionnelle du contrôle et de la mise en œuvre des politiques à travers le jeu des élites sociales", cf. ROUBAN L., "Evaluation et modernisation administrative", in CURAPP, (ed), *L'évaluation dans l'administration*, Paris, Presses universitaires de France, 1993, pp. 34-54.

56. Sur ce point, voir DERLIEN H.-U., RIST R., *op. cit.*, pp. 443-444.

## Perspective : identifier les réformateurs du secteur public

Au regard des tous premiers résultats de notre analyse comparative, il semble évident que le cheminement institutionnel de l'évaluation des politiques publiques ne saurait être retracé et compris sans analyser en profondeur le jeu des acteurs concernés par le développement théorique et méthodologique, l'application concrète et les (anticipations des) effets politiques induits par les résultats mêmes de l'évaluation. Ainsi, et comme nous l'avons montré en détail pour les cas de la Belgique, de la France, des Pays-Bas et de la Suisse<sup>57</sup>, l'ancrage institutionnel, les usages et la maturité de la pratique évaluative dépendent en premier lieu des intérêts défendus, des ressources mobilisées ainsi que des règles et arènes décisionnelles activées par les commanditaires politico-administratifs de l'évaluation (parlement, gouvernement, administration). Ils sont aussi tributaires de l'appropriation de ce nouvel instrument par les organes de contrôle traditionnels (Cour des comptes et commissions parlementaires par exemple) ou par d'autres instances centralisées de pilotage des politiques publiques (Bureau du Plan par exemple) et, finalement, de l'existence d'une communauté d'évaluateurs (académiques, consultants privés et fonctionnaires spécialisés).

De plus, il est possible d'identifier clairement des porteurs de projets issus de ces trois milieux qui, à titre de véritables "entrepreneurs politiques", diffusent le paradigme de l'évaluation et incarnent une condition nécessaire à son institutionnalisation. L'explication première des différentes voies nationales suivies repose en conséquence sur le poids respectif de réformateurs qui, comme le note de manière très convaincante Ph. Bezes<sup>58</sup>, croient au mythe de l'évaluation, dans le sens où, d'une part, ils adhèrent aux idées de modernisation du secteur public que sous-tend l'évaluation des politiques (aspiration réformatrice) et, d'autre part, jugent possibles sa mise en œuvre et sa pérennisation (faisabilité politique). Pareille approche présuppose par contre une étude fine des motivations et du rôle des différents acteurs, lui-même dépendant de leurs dispositions individuelles, attaches organisationnelles et positions institutionnelles.

---

57. JACOB S., VARONE F., *op. cit.*, 2003.

58. BEZES Ph., "Les hauts fonctionnaires croient-ils à leurs mythes ? L'apport des approches cognitives à l'analyse des engagements dans les politiques de réforme de l'Etat. Quelques exemples français (1988-1997)", *Revue française de science politique*, 2000, vol. 50, n° 2, pp. 307-332.