

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

L'ÉVALUATION UN CYCLE DE VIE À LA FRANCAISE

Jean-Claude THOENIG

Jean-Claude THOENIG est directeur de recherche (émérite) au CNRS à Dauphine
Recherche en Management UMR 7088 (CNRS et université Paris Dauphine)
Adresse : jean-claude.thoenig@dauphine.fr

.

L'évaluation, qui est un jugement basé sur des analyses rigoureusement fondées, ne se confond pas avec la science des résultats parfaits. Elle fournit une source de connaissance parmi d'autres références, comme le sens commun ou l'expérience, que les décideurs publics prennent en compte. Les évaluateurs ne sont pas les propriétaires et destinataires de l'évaluation. Mener des politiques publiques, les élaborer et les conduire, trace fondamentalement l'essence du politique : la décision publique n'est pas gouvernée par la science.

L'expérience française d'évaluation couvre trois décennies. Elle est nourrie des évidences mentionnées ci-dessus. Néanmoins elle n'aura pas été exempte de relâchements du bon sens. En même temps elle aura fait preuve d'intelligence pragmatique. Si son histoire reste à écrire, le regard qu'elle mérite suggère de la considération et un peu d'ironie bienveillante.

1. Le temps des bricoleurs

La référence américaine et l'analyse des politiques publiques alimentent la découverte de l'évaluation qui s'opère en France (Cazes 1974; Thoenig 1976). Les pratiques modernes et explicites d'évaluation appliquées à des politiques gouvernementales s'y implantent dès les années 1970.

Elles sont le produit d'initiatives non médiatisées, prises de manière plus ou moins disjointe, traitant de dossiers peu sensibles politiquement, et assumées par des praticiens n'appartenant pas nécessairement à l'élite de l'élite administrative (Deleau 1986). Elles sont conduites dans le cadre d'administrations d'Etat, avec l'aide de quelques chercheurs. (Nioche et Poinard 1984). Une revue comme *Politiques et Management Public*, créée en 1984, joue un rôle privilégié de promotion.

La puissance publique est le destinataire quasi exclusif. L'évaluation n'est pas appropriée par des entreprises politiques et des lobbies pour servir de levier d'argumentation et de débat sur des causes spécifiques ou des intérêts ciblés. Il n'existe en France ni *think tanks* indépendants sur le modèle de la *Brookings Institution* ni groupes de pression de type *Sierra Club*. La contradiction publique basée sur des informations raisonnées n'est pas un savoir-faire établi, pire, une mauvaise manière.

De 1971 à 1978, 36 évaluations dignes de ce label sont menées. De 1978 à 1983, le chiffre passe à 153 (Nioche et Poinard 1985). Les domaines couverts sont les politiques nationales, notamment agricoles, environnementales, de santé publique, d'emploi ou encore d'aide à l'innovation. Quelques colloques aidant,

la référence à l'évaluation devient de bon ton. Des corps d'inspection générale ambitionnent d'ajouter à leur fonction de contrôle de conformité procédurale une compétence d'évaluation des politiques (Gibert et Andrault 1984). La Cour des Comptes s'équipe en doctrine et en savoir-faire (Club Cambon 1988). De l'évaluation, on attend tout ou n'importe quoi, la clé de la réforme administrative ou encore la mise sous contrôle par les ministères de Paris de services extérieurs et d'organismes déconcentrés trop tentés par l'autonomie et par l'opacité. Chacun se sent une vocation d'évaluateur, même s'il ne possède qu'un rudimentaire bagage méthodologique.

La préconisation de dispositifs dits participatifs ou pluralistes devient un outil de marketing (Monnier 1987). En même temps, la tonalité dominante est celle d'une conception pragmatique et opportuniste. L'idéologie du New Public Management reste étrangère à la volonté d'un management public plus rationnel dont l'évaluation participe. Le pays de la technocratie considère avec recul les querelles entre écoles de pensée telles qu'elles se manifestent outre-Atlantique dans les revues professionnelles. L'évaluation affirme moins le règne des quantitativistes bornés que la reconnaissance des sciences sociales molles comme sources de données pertinentes pour l'action publique.

2. La révolution institutionnelle de 1990

Le 22 janvier 1990, le décret que publie le Journal Officiel fait sortir l'évaluation de la pénombre. Le gouvernement de Michel Rocard fonde *ex nihilo* une machinerie installée au sommet de l'Etat et dédiée à la seule évaluation des politiques : un organe de pilotage politique, le Comité interministériel de l'évaluation, un Fond national pour financer les projets, un Conseil scientifique pour garantir la qualité et l'indépendance de l'évaluation. L'annonce fait sensation, compte tenu de la solennité donnée par le Premier ministre à la fois au dispositif et aux objectifs assignés. Pour certains zélotes, son héroïque audace marque la victoire de l'évaluation en France.

L'évaluation est étatisée. C'est d'en haut de l'Etat et par un quasi-monopole que son expansion sera régulée. Un organe unique assure à la fois la supervision des évaluations conduites par des prestataires à façon sur une série de politiques d'Etat et une fonction d'animation - diffusion de la bonne parole à travers le pays. La méthodologie établit sa compétence distinctive et légitime sinon la quasi-tutelle de fait dont il est investi. L'institutionnalisation de l'expertise est poussée à son apogée.

De là à penser que le centre de l'Etat laminera les initiatives qui lui échappent et qu'une économie administrée de l'évaluation se répande, faute de contre-pouvoirs (associations professionnelles, etc), le pas est franchissable par les plus

sceptiques. Par ailleurs, il est possible de voir dans l'évaluation une machine qui, à l'instar de la RCB (rationalisation des choix budgétaires) vingt ans plus tôt, renforce la capacité d'ingérence dans les choix dont disposerait le ministère du Budget par rapport aux ministères dits dépensiers (Thoenig 1971).

La sacralisation opérée en 1990 rend l'évaluation étroitement dépendante de l'autorité gouvernementale. Le dispositif est rattaché directement au Commissariat général du plan. Il destine ses travaux au gouvernement qui lui passe commande. L'évaluation est donc apparemment intégrée à la préparation de la décision politique. La contrepartie est une extrême dépendance à l'attention que porte le pouvoir exécutif aux évaluateurs, surtout lorsque leurs travaux prennent deux sinon trois années à déboucher. Une fois Michel Rocard remplacé à Matignon en 1992, le dispositif sombre rapidement dans la zone d'indifférence, le creux étant atteint entre 1995 et 1997 sous le gouvernement dirigé par Alain Juppé. Lionel Jospin, par une circulaire du 28 décembre 1998, relancera le dispositif. (voir www.evaluation.gouv.fr)

Imaginer que le décideur soit si proche de l'évaluateur paraît naïf dans un pays dont quelques grands corps trustent les fonctions dirigeantes. Voir débouler des experts autodidactes car non issus de l'ENA et de Polytechnique, mobiliser les sciences sociales, sont des initiatives qui composent un scénario improbable. Les faits de l'expérience suggèrent pourtant que la *policy evaluation* et la *policy analysis* composent deux approches rarement jumelées, que l'évaluation est un produit en soi, non associé par nature à la décision. Qu'ensuite quelque décideur en fasse quelque chose et en tire usage est une tout autre affaire.

De surcroît l'évaluation est lestée d'ambitions sociétales. Outre l'éclairage plus rationnel du pouvoir exécutif sur de grands dossiers ou considérés comme tels, le décret de 1990 et ses inspireurs prétendent faire d'une méthodologie d'étude un levier de renouvellement du débat démocratique. L'évaluation sera transparente. Les travaux des experts seront intégralement publiés. Elle sera aussi et surtout participative et pluraliste. Une instance spécifique par politique sera constituée qui conduira le processus depuis la définition des buts et des méthodes à satisfaire, le choix des producteurs à qui seront sous-traitées les études, jusqu'à la validation des résultats ainsi fournis, leur interprétation et l'écriture du rapport final. Or sa composition devra doser les membres nommés pour leur expertise et ceux qui représenteront des divers *stakeholders* concernés (ministères et collectivités locales, groupes d'intérêts socio-économiques, associations défendant des causes morales, etc). Débat et rigueur sont les deux mamelles du jugement expert basé sur la connaissance.

Une contradiction émerge de l'observation des pratiques internationales. Comme le suggère l'exemple de la Banque Mondiale, les évaluateurs comptent

d'autant plus que leur service est totalement coupé du reste de l'organisation. Par ailleurs un modèle participatif induit des conflits de pouvoir, rallonge les délais et favorise les non-dits sinon les compromis peu innocents.

3. Le verre à moitié vide est pourtant aussi à moitié plein

L'opinion des cercles avertis est souvent cyclothymique quand il s'agit de juger des avancées de la modernisation administrative. Tel est le cas pour l'évaluation en France. A l'euphorie sans limites succède dès le milieu des années 1990 un scepticisme exagéré. Il est de bon ton d'en parler désormais comme d'un phénomène du passé, d'en regretter le prétendu déclin et d'en accabler les illusions perdues (Hespel 2002)

Le fait est que l'évaluation est entrée dans une phase de désenchantement. Elle n'agit pas comme la potion magique de la modernisation rationnelle. Elle ne tombe pas non plus aux oubliettes. Plus simplement, elle sert à qui veut s'en servir (Chanut 2002)

Est-il vrai que le microcosme des évaluateurs parle d'autant plus d'évaluation qu'il en fait peu? Une observation pondérée des pratiques et des usages indique que la France se classe fort honorablement par rapport aux autres pays de l'OCDE. Au total environ 1.500 évaluations auront été menées à terme à ce jour, même si certaines ressemblent à des quasi-évaluations, plus proches de l'audit ou de l'étude que du type-idéal de l'évaluation qu'enseignent les manuels. Une assez large palette de méthodes et d'approches aura été mobilisée. La plupart des domaines de l'intervention publique auront été explorés. Les commanditaires auront été fort divers : le gouvernement, le parlement, les ministères, les collectivités locales, des agences publiques autonomes. La liste est impressionnante des institutions publiques qui, en tout cas facialement, revendiquent un recours à l'évaluation. Pour celles établies au niveau national, on peut recenser plus de quarante d'entre elles. Une diffusion assez vaste et différenciée sera advenue à travers le pays, souvent sans effet d'annonce et sans publicité. Le recours à l'évaluation aura été la plupart du temps volontaire, même si, dans quelques cas, il résulte d'une obligation procédurale (par exemple pour les fonds structurels abondés par l'Union Européenne et les contrats de plan entre l'Etat et les régions) sinon de la loi.

Même le dispositif interministériel reprend en 1999 quelque vigueur momentanée, malgré la défaillance de ses membres et l'indifférence du Commissariat du plan. Entre 1991 et 2002, il traite 42 évaluations. Les plus connues sinon influentes portent sur la contractualisation des relations entre l'Etat et les universités publiques, la lutte contre l'alcoolisme et le tabagisme, ou encore la préservation des zones humides. Les travaux menés de la mise en

œuvre des règles de sécurité routière inspirent en 2003 le renforcement des mesures de lutte contre le fléau imposé par la Présidence de la République. Néanmoins la défaillance de ses membres et l'indifférence du Commissariat général du Plan vont remettre en sommeil le dispositif au plus vite.

Quelques traits caractérisent les usages et les pratiques en France.

L'évaluation porte principalement sur l'efficacité des politiques. Elle cerne des effets ou impacts extérieurs de l'action publique, voulus ou induits. La mise en œuvre notamment retient l'attention dans la mesure où elle révèle des dysfonctions sur le terrain qui pervertiraient les objectifs visés ou annoncés. En revanche, l'efficacité, l'aspect *value for money*, est nettement moins prise en compte, sinon considéré comme trivial. A cela des raisons diverses allant d'un équipement comptable et budgétaire encore imprécis à une pression politique n'affichant guère les arguments de productivité et de qualité dans la concurrence électorale ou pour légitimer les producteurs de biens publics. Par ailleurs la France n'est pas la Grande-Bretagne : l'agnosticisme à l'égard de qui, du privé ou de l'Etat, fournit un service dit public n'y est pas vraiment développé. Elle n'est pas non plus la Nouvelle-Zélande du début des années 1980 : ses politiciens marquent un peu plus de retenue quant à l'exploitation effrénée de la faillite de l'Etat et de ses finances.

L'évaluation marque aussi une émergence significative du recours aux sciences sociales. Des langages et techniques ont acquis un droit de cité de la part des commanditaires publics : la sociologie des organisations, les démarches inductives, l'entretien en profondeur, l'attention portée aux opinions, etc. Le culte du chiffre se restreint, même si par ailleurs ce sont plutôt des statisticiens qui occupent les postes de responsabilité dans les services publics d'évaluation de l'Etat. Le recours aux données molles est parfois si laxiste qu'il devient difficile de qualifier si une évaluation relève d'autre chose qu'une simple monographie impressionniste.

La production d'évaluation reste assez fortement internalisée par le secteur public, même si on constate un recours à des évaluateurs et sous-traitants venant du privé, exceptionnellement du monde académique. Les commanditaires français préfèrent la hiérarchie au marché. Les services et autres dispositifs se sont multipliés qui, à temps partiel ou par compétence exclusive, passent commande, produisent ou supervisent des gestes évaluatifs. Ainsi les ministères de l'Etat ont mis en place des capacités propres - l'Equipement, les Affaires Sociales, la Santé, l'Agriculture, l'Education Nationale -, certains dès les années 1980. Ces dispositifs plus moins solides et crédibles permettent de sanctuariser leur domaine respectif. Quelques régions et départements se lancent à leur tour dans une institutionnalisation plus ou moins musclée. L'avantage

relatif par rapport au recours au marché est de donner une visibilité à l'évaluation sinon d'assurer une certaine continuité et d'espérer échapper aux aléas ou aux lubies des décideurs. L'évaluation se banalise et est dédramatisée. Elle est accessible à un grand nombre. Il n'est pourtant pas certain qu'une véritable culture de l'évaluation s'installe. En aval, l'internalisation ralentit l'émergence d'un marché de professionnels dédiés, milieu qui remplisse une fonction de référence en matière de bonnes pratiques. Les services d'études permanents et les instances pluralistes *ad hoc* abaissent la barrière d'entrée à la qualification d'évaluateur. En amont, l'internalisation ne garantit pas que le décideur politique accorde une attention plus forte à l'évaluation parce que ses producteurs se situent dans sa propre institution.

L'évaluation a fait son trou, même si certains espéraient mieux. Elle a réglé son compte aux *ayatollah* de tous poils, qui espèrent une gouvernance par la vérité des chiffres ou qui rêvent d'une agora évaluatrice (Kessler, Lascoumes, Setbon et Thoenig 1998). Elle a pris sa place dans le rang des techniques et démarches qui composent la palette du management public, plus précisément pour ce qui concerne la connaissance des conditions de management des politiques publiques. L'échec du rêve de 1990 aura masqué un bilan qui, tout compte fait, à ce stade, n'est pas déshonorant.

Une forme de quasi-évaluation pervertie revient en force : celle des rapports demandés à une personnalité à chaque fois qu'un nouveau gouvernement ou qu'un nouveau responsable politique arrive aux affaires. Choissant la personnalité, rédigeant sa lettre de mission, le prince peut aussi prendre dans le rapport ce qui lui convient. Mieux, cette formule étouffe le débat et la controverse, faute de rigueur méthodologique dans l'analyse.

4. Apprendre à partir des pratiques

Un syndrome habituel des *do-gooders* français en matière de modernisation de l'Etat est de se tourner vers des références étrangères sous prétexte que l'herbe y est plus verte, OCDE et autres instances internationales aidant. Les références américaines et britanniques aux usages réels ou affichés de l'évaluation ont certainement impressionné quelques observateurs français à un moment ou à un autre. Mais, à la différence de l'importation opérée en France à la fin des années 1960 des méthodes développées aux Etats-Unis de *PPBS* et rebaptisées sous le vocable de RCB, le regard porté dès les années 1970 sur l'évaluation s'est montré, sauf exceptions, plus nuancé et l'enthousiasme moins inconditionnel. Il ne s'agissait pas d'imiter pour imiter. C'est curieusement quelques années plus tard et de façon purement endogène que se développera une sorte de contre-modèle breveté indigène de type pluraliste et participatif.

La maturité du cycle de vie en matière d'évaluation se traduit par la capacité d'une expérience nationale à tirer des enseignements à l'écoute de ses pratiques et usages. Des illusions sont tombées et des surprises ont émergé qui ont modelé le pragmatisme à la française. On en mentionnera quatre (Thoenig 2002).

S'il est nécessaire que les commanditaires élaborent un cahier de charges contractuelles, la qualité et la pertinence de l'évaluation ne sont pas pour autant garanties par un cahier détaillé et pointilleux élaboré par des fonctionnaires de la commande publique qui sont labellisés comme étant des évaluateurs. Pour éviter le n'importe quoi, l'expérience française suggère que l'aval – la compétence professionnelle des évaluateurs - compte plus que l'amont – la rationalité du cahier de charges.

S'il est important que les gestes de l'évaluation constituent une compétence distincte, il ne faut pas pour autant espérer trop de l'érection d'un statut professionnel trop certifié de l'évaluateur. Deux conditions doivent être satisfaites : une connaissance du domaine qu'il couvre (santé, éducation, etc.) des qualités analytiques. L'approche par l'efficacité et celle par l'efficacité mobilisent des profils différents à cet égard.

Les évaluations faisant appel à des instances spécifiques offrent une utilité moins par leur caractère de représentativité que par leur levier d'échange. En d'autres termes, participation et pluralisme aident s'ils contribuent à élargir les couloirs cognitifs de leurs membres, leur façon de poser un problème et de donner sens à des informations. La clé dans ce cas réside dans la manière dont est pilotée l'instance, donc dans le choix de ses animateurs et dans le rythme de ses activités. Le phénomène d'*advocacy* ralentit les délais et paralyse l'analyse. Le processus fait la différence : des jugements équitables sont des jugements partagés car co-construits.

Enfin, si les réformes administratives restent une nécessité indiscutable, utiliser l'évaluation comme un levier privilégié pour améliorer le mode de fonctionnement des bureaux et changer les comportements des fonctionnaires se révèle une voie hasardeuse. Mettre en relief des dysfonctionnements qui grèvent la mise en œuvre d'une politique publique remplit une fonction d'alerte pour les décideurs et de sensibilisation plus générale. En espérer *ipso facto* un déclencheur de changement est une illusion. L'évaluation n'est pas un geste de management des personnes et des structures. Il est une source de connaissance parmi d'autres dont des jugements peuvent être tirés qui se fondent sur des informations traitées selon des canons de rigueur acceptables.

Bibliographie

Cazes B. (1974) *L'analyse stratégique*. Paris, Commissariat général du Plan.

Chanut V. (2002) « L'évaluation : affaire d'Etat ou question d'organisation ? », *Politiques et Management Public*, 20 (4) : 1-32

Club Cambon (1988) « L'évaluation des politiques publiques », *Politiques et Management Public*. 2 : 98-117

Deleau M. (dir.) (1986) *Evaluer les politiques publiques : méthode, déontologie, organisation*. Paris, La Documentation Française.

Gibert P. et Andrault M (1984) « Contrôler la gestion ou évaluer les politiques », *Politiques et Management Public*. 2 :123-132

Hespel V. (2002) « L'avenir institutionnel de l'évaluation des politiques publiques », *Politiques et Management Public*, 20 (4) : 51-65

Kessler M., Lascoumes P., Setbon M. et J.-C. Thoenig (dir.) (1998), *L'évaluation des politiques publiques*, Paris, L'Harmattan.

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

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

Nioche J.-P. et Poinsard R.(1985) « Public policy evaluation in France », *Journal of Policy Analysis and Management*. 5, 1 : 58-72

Thoenig J.-C. (1971) « Le PPBS et l'administration publique. Au-delà du changement technique », *Annuaire International de la Fonction Publique*. Paris, Institut international d'administration publique : 97-114

Thoenig J.-C. (1976) « Enseigner le management public », *Enseignement et gestion*,15 : 3-33

Thoenig J.-C. (2004) « L'évaluation en actes : leçons et perspectives », *Politiques et Management Public*, 20 (4) : 33-50