



HAL
open science

La (f)utilité gouvernementale de l'évaluation des politiques publiques, quelques leçons américaines et françaises

Vincent Claude Spenlehauer

► **To cite this version:**

Vincent Claude Spenlehauer. La (f)utilité gouvernementale de l'évaluation des politiques publiques, quelques leçons américaines et françaises. LIEPP Working Paper, 2016, 49. hal-01954150v2

HAL Id: hal-01954150

<https://enpc.hal.science/hal-01954150v2>

Submitted on 13 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License



SciencesPo

LABORATOIRE INTERDISCIPLINAIRE
D'ÉVALUATION DES POLITIQUES PUBLIQUES

LIEPP Working Paper

Mars 2016, n°49

La (f)utilité gouvernementale de l'évaluation des politiques publiques, quelques leçons américaines et françaises

Vincent Spenlehauer

École nationale des Ponts et Chaussées

vincent.spenlehauer@enpc.fr

www.sciencespo.fr/liepp

© 2016 by the authors. All rights reserved.

La (f)utilité* gouvernementale de l'évaluation des politiques publiques, quelques leçons américaines et françaises

Vincent Spenlehauer (École nationale des Ponts et Chaussées)

Introduction

L'évaluation des politiques publiques (EPP) est née aux Etats-Unis dans les années 1960, au croisement d'une offre et d'une demande. L'offre venait des sciences sociales américaines, qui présentent à partir des années 1950 un front relativement uni d'expertise professionnelle pertinente en termes d'analyse de politiques publiques. Quant à la demande, elle est soutenue à la fois par les promoteurs et les détracteurs de l'interventionnisme dont fait preuve l'administration Kennedy, puis Johnson, dans de nombreux domaines d'action publique (santé, logement, éducation, transports, protection de la nature, déségrégation, guerre froide ou chaude, conquête spatiale...)¹. Dès

* Ce titre est inspiré de Gusfield J. 1975, The (F)Utility of Knowledge?: The Relation of Social Science to Public Policy toward Drugs, *Annals of the American Academy of Political and Social Science*, Vol. 417, Drugs and Social Policy (Jan., 1975), pp. 1-15.

¹ Sans compter les éléments de contexte qui ont favorisé le développement industriel de l'EPP comme par exemple la réaffectation à la fin des années 1960 des moyens en personnels du *Program Planning & Budgeting System* fédéral (Wildavsky 1969).

lors, l'ancrage dans les sciences sociales académiques et le projet de transformer de la « bonne connaissance » sur l'action publique en « bon gouvernement » constituent les deux principaux gènes de l'EPP.

Cependant, avec le temps et la diffusion de l'EPP dans les autres pays occidentaux, ces deux gènes ont largement perdu de leur caractère dominant. En 1970, Donald T. Campbell se posait en prophète d'une nation américaine se muant toute entière en « experimenting society » (Campbell 1969 ; Dunn 1998). Désormais, les social scientists, collègues américains inclus, sont des acteurs minoritaires de l'activité évaluative. C'est le cas en France, comme le révèle l'examen de la liste des études financées par le Fonds d'Expérimentation pour la Jeunesse, ou de la façon dont ont été construites en 2012 et 2013 les EPP décidées en Comité interministériel de modernisation de l'action publique. L'ambition de réaliser des EPP utiles sur un plan gouvernemental a pris elle-aussi un tour récessif. Les praticiens reconnaissent de façon unanime la très fréquente futilité gouvernementale des EPP (Vedung 2000 ; Flyvbjerg 2001 ; Chanut & Cousquer 2003 ; Nutley et al. 2007 ; Mouterde 2009).

On observe donc d'un côté une théorie évaluative qui continue de porter principalement sur la question méthodologique et, de l'autre, des praticiens qui prospèrent, même en France (Matyjasik 2010 ; Fouquet 2014). Et, corollaire à cela, le peu d'impact réformateur des EPP sur leur objet ne semble pas contrarier grand monde, même si, à l'origine, l'EPP avait été inventée pour cela.

La présente contribution rassemble un certain nombre de résultats d'études et d'observations, américaines en majorité, concernant l'impact gouvernemental des EPP. L'EPP est ici définie comme le processus de production de connaissances qui :

- 1) est officiel, c'est-à-dire publiquement et légitimement commandité par une ou plusieurs autorités publiques ;

- 2) porte sur une action gouvernementale identifiable, c'est-à-dire justifiée politiquement (le pourquoi) et qualifiée administrativement (le comment)² ;
- 3) s'appuie sur un protocole de production de connaissances explicité (sous la forme en général d'un cahier des charges avec annexes méthodologiques) ;
- 4) et, enfin, place l'efficacité de l'action et les possibilités d'amélioration au cœur de son questionnement³.

L'approche scientifique à l'utilité conditionnelle

Les premières études sur l'utilité gouvernementale des EPP datent de la fin des années 1960 aux Etats-Unis (Weiss 1967). Elles portent la marque du scientisme de l'époque dans la mesure où l'EPP se conçoit comme un processus en trois phases bien distinctes : 1) une décision politique lançant une évaluation, 2) laquelle est organisée et réalisée par des scientifiques 3) lesquels livrent les résultats de l'évaluation aux acteurs politiques qui en usent de diverses manières pour faire évoluer la politique évaluée. Un des avantages de cette approche réside dans sa simplicité et donc son applicabilité à de nombreux cas réels d'EPP.

Parmi les leçons les plus saillantes de ces études américaines, un premier enseignement concerne la décision politique de lancer une EPP. Celle-ci s'inscrit rarement dans un schéma rationnel instrumental simple. Elle procède en général d'un accord entre de nombreuses parties prenantes poursuivant des objectifs différents et parfois contradictoires. Certains acteurs attendent de l'évaluation qu'elle les aide à imposer une option

² De ce point de vue, une expérimentation constitue un dipôle formé par une action gouvernementale à petite échelle et son évaluation.

³ On pourrait ajouter que l'EPP a aussi vocation à alimenter le débat public, ce qui lui conférerait une utilité démocratique.

innovante de politique publique. D'autres acteurs envisagent la même EPP comme une manœuvre de retardement, voire, comme un lot de consolation symbolique à offrir aux innovateurs et facilitant en conséquence le maintien du statu quo (Leviton & Edward 1981). Il existe encore d'autres d'usages de l'EPP, à des fins de clarification, comme instrument politique permettant de débloquent une situation, etc.

La seconde phase de l'EPP, animée essentiellement par les scientifiques, est caractérisée par la question de sa crédibilité interne (ou scientifique) et externe (ou politique). En ce qui concerne la première, l'histoire maintenant pluri-décennale de l'EPP est jalonnée de tentatives visant à imposer au reste de la communauté scientifique un étalon-or méthodologique, le dernier en date étant le paradigme expérimentaliste sur base d'essais randomisés contrôlés (Vedung 2010). Cependant, l'expérience montre sans conteste qu'en raison de la grande diversité des politiques publiques et de la complexité de chacune d'entre elles, une évaluation scientifiquement solide ne peut être que composite sur le plan méthodologique (Barbier 2014 ; Trochim 2009).

Il ne s'agit pas ici de prôner un anarchisme de méthode mais plutôt de mettre en avant la valeur d'une approche méthodologique pluraliste et ad hoc. Par exemple, dans le cas d'une EPP ex post portant sur un dispositif d'action fonctionnant depuis deux ans, l'étude a posteriori de ses effets de court-terme va probablement reposer sur des méthodes assez inductives, alors que l'appréciation ex ante des effets de long terme requiert plutôt une approche hypothético-déductive. Par ailleurs, ces deux corps méthodologiques ne ressembleront pas à celui nécessaire à la compréhension fine des facteurs endogènes et exogènes à la politique évaluée pouvant expliquer les niveaux d'efficacité observés.

Dans cette perspective d'une inévitable pluralité méthodologique de l'EPP, la notion de crédibilité interne n'a plus guère de sens. Dès lors que le

regard embrasse tout le spectre méthodologique, c'est la question de la crédibilité externe qui s'impose. La manière plus ou moins harmonieuse avec laquelle différents éléments méthodologiques s'agencent, se complètent et interagissent en bonne intelligence constitue une des principales variables dont dépend la crédibilité d'une EPP. Une évaluation dont les animateurs scientifiques offrent le triste spectacle de querelles interdisciplinaires et/ou paroissiales s'expose au risque de la déconsidération et donc de l'inutilité (Sabatier 1978)⁴.

En ce qui concerne la phase d'utilisation des résultats de l'EPP, de nombreuses études conduites aux Etats-Unis dans les années 1980⁵ convergent sur au moins deux points. Premièrement, ces études montrent que la manière dont les chercheurs-évaluateurs communiquent les résultats de leurs EPP et soignent leur rhétorique de valorisation extrascientifique peut faire une grande différence (Mandell 1989). Dans un *information-rich world* (Simon 1997) une EPP qui ne parvient pas à capter l'attention de possibles utilisateurs, quand bien même ceux-ci ont participé à la commande, a toute chance de rester lettre morte. Un des grands défis tient par exemple à la synchronisation du déroulement de l'EPP et de la vie de la politique évaluée elle-même. Deuxièmement, les traits professionnels des destinataires d'une EPP conditionnent beaucoup la réceptivité de ces derniers. L'enquête conduite à ce propos par le politiste Barry Bozeman (1986) en apporte une démonstration sans équivoque : ainsi le quantitativisme prononcé d'une EPP intéressera l'homme politique professeur de médecine dans le civil et rebutera son collègue avocat au

⁴ Dans le même esprit, mais bien plus tard, Paul Sabatier proposera un modèle explicatif du changement non-incrémental dans les politiques publiques : *l'advocacy coalition framework* (Sabatier 1988). Autrement dit : s'ils veulent avoir une influence notable sur la politique qu'ils évaluent, les chercheurs-évaluateurs ont plutôt intérêt à se coaliser.

⁵ Dès son arrivée au pouvoir en 1981, Ronald Reagan a opéré d'importantes coupes budgétaires dans les secteurs d'intervention publique les plus demandeurs d'évaluation (emploi, logement, éducation, environnement...), obligeant sans le vouloir la profession évaluative à s'interroger sur son utilité (GAO 1987).

barreau; ce qui rejoint les considérations précédentes concernant la pluralité méthodologique des EPP.

Encadré 1

Illustration : le cas du Plan Pandémie Grippale et la nécessaire pluralité méthodologique de l'évaluation

On se place dans le cas hypothétique de l'évaluation de la politique publique sanitaire intermittente appelée « Plan Pandémie Grippale ». Il s'agit d'examiner son déclenchement et sa mise en œuvre entre juillet 2009 et janvier 2010 à la suite de l'alerte pandémique mondiale du virus H1N1 lancée par l'OMS en juin 2009. Deux grandes questions s'imposent :

- 1) comment expliquer le faible taux de vaccination (9%) enregistré, sachant qu'un tel taux va de pair avec une efficacité quasi-nulle pour juguler ce qui semblait être, au départ, une pandémie ?
- 2) comment transformer la politique évaluée pour élever le taux de vaccination à au moins 60% si la France est dans le futur confrontée à une vraie pandémie grippale ? Le cas échéant, quels sont les autres registres d'action que celui de la vaccination ?

Ces deux questions sont emboîtées. Les ramifications méthodologiques de la première question sont déjà nombreuses. Quel est le gradient de vaccination selon les zones, le sexe, l'âge, la CSP, les lieux de travail ? Observe-t-on un fort taux de vaccination parmi les personnels hospitaliers dans les centres accueillant des personnes atteintes du virus ?

Comment fonctionne l'OMS et pourquoi a-t-elle maintenu son niveau d'alerte maximal, alors que dès la fin août 2009 des analyses rassurantes sur la virulence de H1N1 parvenaient de l'hémisphère sud (Crouzel et al. 2015) ?

Comment se sont comportés les laboratoires pharmaceutiques ? Peut-on clairement modéliser le renforcement de leur position oligopolistique à l'occasion de la crise ?

Que peuvent nous apprendre des regards comparatifs jetés sur des pays aux taux de vaccination similaires (8% en Grande-Bretagne, 10% en Allemagne) ou différents (44% au Canada, 65% en Suède) ?

Quel est le poids de la variable « *politics of blame avoidance* » (Weaver 1986) dans ces différents pays (il est probablement fort en France depuis l'affaire du sang contaminé) ?

Les comparaisons diachroniques sont également possibles, en particulier avec la « grippe de Hong-Kong » de 1968-69 à laquelle on impute environ 30 000 décès rien qu'en France.

Certaines questions « uchroniques » sont aussi pertinentes ; ainsi, la gestion du H1N1 aurait-elle emprunté un autre chemin si le ministère de la santé n'avait pas été au même moment en pleine négociation, dure, avec les médecins de ville ? Etc.

Carol Hirschon Weiss, la fondatrice des recherches en matière d'utilisation d'EPP, attire notre attention sur le fait que l'absence d'impact substantiel de court terme des EPP ne doit occulter ni leur impact stratégique de moyen-terme, ni leur impact conceptuel de long-terme (Weiss 1979). Weiss propose le modèle des trois « I » - Idéologies, Intérêts, et Informations - pour qualifier l'environnement de la politique considérée et l'écho que peut y trouver une EPP (Weiss 1983). Selon cette approche macrosociologique, l'influence d'une EPP dépend largement de la plus ou moins grande congruence globale de ses résultats avec le positionnement des acteurs principaux de la politique évaluée – c'est à dire avec le *policy subsystem* (Howlett & Ramesh 1995)- sur les trois « I-dimensions » mentionnées.

Est-ce à dire que l'évaluation de toute politique publique procédant d'un arrangement global et stable entre acteurs sur ces trois dimensions est condamnée à l'inutilité ? Le message de Weiss et de son modèle penche peut-être davantage vers le réalisme que vers le défaitisme. Weiss indique d'abord qu'une EPP réalisée n'atterrit jamais sur une page blanche. Les résultats de l'EPP s'insèrent dans des systèmes de pouvoirs animés par des acteurs qui portent des valeurs. Ainsi, les conclusions d'une évaluation concernant la politique de bridage des moteurs à des fins de sécurité routière est susceptible d'interagir avec les valeurs défendues par les acteurs, comme par exemple celle de la responsabilité du conducteur (d'où la maxime bien connue : « *cars don't kill people, people kill people* »). Ces acteurs défendent aussi des intérêts, par exemple celui des individus travaillant dans le secteur automobile qui pourraient être menacés de chômage par la mise en place de cette politique. Ces acteurs, enfin, disposent souvent d'informations autres que celles issues de l'EPP - par exemple, le fait que de façon générale, les constructeurs automobiles servent davantage la sécurité routière qu'ils ne la desservent. A partir de ce postulat, Weiss suggère qu'une EPP est d'autant plus influente que son objet a été circonscrit en amont au sein d'espaces incertains de la politique

à évaluer, c'est à dire au sein d'espaces encore non verrouillés autour d'idéologies et d'intérêts très structurés, et dans lesquels l'information est encore lacunaire et/ou controversée.

Weiss invite en conséquence les *evaluation researchers* à ne pas se soucier uniquement de la scientificité de leurs dispositifs évaluatifs. Il leur faut aussi procéder, avant construction, à une topologie politique du domaine d'action publique au sein duquel l'évaluation va tenter d'inoculer le virus du changement rationnel. En ce sens, sans se dépendre totalement de son scientisme d'origine, Weiss ne dénie pas toute pertinence à l'approche postmoderne de l'EPP qui voit cette dernière comme un entrelacement serré de jeux de pouvoirs et de processus de productions de connaissances.

L'approche postmoderne et interactionniste

L'œuvre entière de Charles Lindblom s'apparente à un long théorème de l'impossibilité de rationaliser les politiques publiques via leur évaluation organisée (Lindblom 1990). Pour Lindblom, la plupart des décideurs politiques se caractérisent non seulement par une rationalité limitée mais aussi par la conscience d'incertitudes caractérisant 1) les objectifs à atteindre 2) les moyens nécessaires et/ou à disposition pour atteindre les objectifs 3) les conséquences des décisions couplant des objectifs et des moyens. Dans ce modèle, les acteurs négocient donc pied à pied les objectifs, les moyens et les conséquences possibles, ce qui conduit en général à des décisions mesurées, s'écartant peu des sentiers battus, mais satisfaisant une majorité de parties prenantes ou au moins ne suscitant pas de véto chez elles.

Cette façon de faire est tellement pétrie de prudence qu'elle n'engage pas les prochaines décisions à prendre, d'où son appellation d'incrémentalisme disjoint. Lindblom et Braybrooke (1963) confirment leur défiance à

l'encontre de l'idée d'une possible rationalisation du processus de production des politiques publiques par les sciences sociales. Celles-ci produisent des connaissances qui sont largement relativisées ou diluées lors de leur réception par les praticiens de l'action publique. En ce sens, Carol Weiss partage la conception sceptique de Lindblom et l'illustre avec l'exemple suivant dans une communication de 1973 (1993): « *If a decision-maker thinks it is important for job trainees to get and hold on to skilled jobs, he will take negative evaluation findings seriously, but if he is satisfied that job training programs seem to keep the ghettos quiet, then job outcome data mean much less* » (Weiss 1993, p.99).

Cependant, d'autres lecteurs de Lindblom s'en sont inspirés tout autrement, pour promouvoir l'idée d'une imbrication forte entre science et politique tout au long d'un processus d'EPP (ex. Patton 1982, 1997 (1978) ; Guba & Lincoln 1989 ; Huberman 1989). Selon Michael Quinn Patton, l'*evaluation researcher* est par principe responsable de maximiser les chances d'utilisation de l'EPP qu'il conduit (Patton 1988). Assumer cette responsabilité implique pour le chercheur-évaluateur de travailler avec l'assemblée des principaux « détenteurs d'enjeux » (*stakeholders*)⁶ de la politique à évaluer. Cette assemblée ou « instance pluraliste d'évaluation » (IPE) va gérer le processus d'évaluation du début à la fin, de la circonscription de l'objet d'étude et la conception du cahier des charges à la restitution des résultats, en passant par toutes les petites décisions de réorientation évaluative chemin faisant.

Eric Monnier (1987) résume les trois rôles que doivent remplir les chercheurs-évaluateurs dans la vision « constitutionnaliste » de l'EPP de Patton : i) méthodologue ii) maïeuticien pour chaque protagoniste qui doit « accoucher » de ses théories d'action et objectifs poursuivis iii) médiateur

⁶ Plus l'assemblée est large plus l'évaluation devient « *participative* ». Ce type d'évaluation est d'autant plus indiqué que les dispositions de la politique à évaluer font l'objet d'importantes actions collectives, non-recours inclus (Warin 2014).

entre parties prenantes, en facilitant les « *partisan mutual adjustments* » dont parle Lindblom dans son livre de 1965. Patton tranche le nœud gordien du chercheur-évaluateur, en incitant les évaluateurs à rendre leurs évaluations utiles par construction, en créant les conditions d'une interaction intelligente avec les utilisateurs potentiels constitués en assemblée. Dans la perspective postmoderne de l'*utilization-focused evaluation* (UFE) - où le « scientifique » cogère la conduite de l'EPP avec le « profane » - c'est dans le processus d'interaction et non au moment de la livraison de rapports écrits, intermédiaires ou finaux, que réside l'utilité d'une EPP. Cette vision présente deux autres avantages par rapport à l'EPP moderne. D'abord, les praticiens d'une politique publique à évaluer détiennent de l'information, à la fois personnellement et à travers les clés institutionnelles de corpus d'information auxquels il importe d'accéder. Il est donc impératif d'interagir avec eux de façon cadrée pour produire une EPP cognitivement solide et crédible. Ensuite, l'interaction avec les praticiens permet de dépasser le stade de l'évaluation seulement récapitulative, pour accéder à celui de « *formative evaluation* » (Scriven 1967), autrement dit, l'évaluation réformatrice de la politique évaluée.

Cependant, deux difficultés subsistent. Le risque de « gauchissement » (traduction du terme « *impairment* » employé par Lindblom (1990)) et d'instrumentalisation de la scientificité d'une EPP, en raison du tour asymétrique que prend souvent l'interactivité entre scientifiques et politiques, est important. Le cas assez récent de l'évaluation expérimentale du Revenu de Solidarité Active constitue un exemple éloquent (Barbier 2012 ; Bouchikhi et al. 2011). A l'automne 2007, quelques mois après avoir été convaincu par le nouveau Haut Commissaire aux Solidarités Actives contre la Pauvreté qu'ils pourraient conduire une évaluation robuste sur la base d'essais randomisés contrôlés sur une durée de trois ans (comme esquissé dans le PLF2006), le consortium d'évaluateurs attirés

(économistes pour la plupart) comprend plusieurs choses⁷ : 1) pour cause de calendrier parlementaire contraint, la durée de l'expérimentation n'excèdera pas huit mois, d'où une chute de sa significativité statistique ; 2) les Conseils généraux et « leurs » CAF se chargeront eux-mêmes de circonscrire les zones tests et témoins ; 3) celles-ci seront délibérément choisies là où la conjoncture locale d'emploi paraît plutôt dynamique, l'argument discutable étant qu'afin de savoir si le RSA supprime bien les « dés-incitations » à l'emploi du RMI il faut que des emplois à pourvoir existent ; 4) la question du taux important de non recours au « RSA activité » (estimé en 2011 à environ 2/3 des éligibles) n'est pas prioritaire pour l'expérimentation, l'argument officiel étant que le retour à l'emploi des RMistes prime sur le soutien aux travailleurs pauvres non-Rmististes puisque l'on compare RSA et RMI⁸. Ainsi, en acceptant le principe d'une co-construction interactive d'une EPP ayant coûté plus de 40 millions d'euros, les scientifiques voient leurs ambitions scientifiques se faire déstructurer par les idéologies, les intérêts, les informations préexistantes des détenteurs d'enjeux du passage du RMI au RSA. Une question mérite alors d'être posée : pourquoi la plupart des chercheurs impliqués sont-ils restés dans le navire alors que de multiples biais le dévoyaient scientifiquement⁹ ?

⁷ En l'espèce, l'instance d'évaluation s'appelle le « Comité d'évaluation des expérimentations » et est composée de trois collèges de dix personnes chacun : celui des départements expérimentaux, celui des représentants de l'Etat et des organismes payeurs (Budget, Trésor, DARES, DGCL, CNAF...) et celui des personnalités qualifiées (donc sept académiques).

⁸ Officieusement, il existe un pacte tacite avec le Budget pour que le passage RMI>RSA ne creuse pas trop le trou des prestations sociales (un taux nul de non-recours au RSA activité signifierait le versement de trois ou quatre milliards d'€ de prestations en plus par an). Ce pacte est tacite car illégal puisque que le RSA, comme le RMI est un droit. De fait, c'est après la généralisation du RSA en 2009 que le Comité national d'évaluation du RSA a été autorisé à traiter sérieusement de la question des travailleurs pauvres ne bénéficiant pas du RSA activité.

⁹ Une réponse à cette question est apportée plus loin dans le texte.

Un autre type d'ingénuité politique peut également être reproché au paradigme de l'UFE. La composition soignée d'une instance pluraliste d'évaluation ne suffit pas à ce que celle-ci fonctionne structurellement comme le creuset d'une réforme partagée et raisonnée de l'action évaluée. Une telle pensée revient à ignorer le fait qu'une part importante du travail de composition est de facto réalisé au gré d'interactions complexes et discrètes entre l'institution chef de file et les institutions sollicitées pour y être représentées ou insistant pour l'être (Frederiks et al. 2002). Cela revient à oublier aussi que la plupart du temps, les institutions représentées dans une IPE y dépêchent des personnes moins pour s'engager en leur nom devant les autres que pour observer et rapporter les informations nécessaires afin de parer à toute éventualité. Ainsi, c'est autant, sinon davantage la gestion ad hoc de l'IPE que le soin apporté à sa composition qui va déterminer l'utilité gouvernementale de l'EPP.

Dans ces conditions, est-il encore possible de réaliser des évaluations qui soient de bonne tenue scientifique et contribuent directement à l'amélioration des politiques évaluées ?

Beaucoup d'EPP sont utiles dans la mesure où elles remplissent diverses fonctions d'utilité portées par divers acteurs pris séparément les uns des autres. La question est donc de savoir si l'on peut combiner et agréger ces différentes fonctions d'utilité, de façon à produire des EPP globalement utilisables pour améliorer les politiques publiques.

L'analyse du champ des utilisations d'une EPP doit dépasser le domaine de la politique évaluée. Si une société de conseil rémunérée pour procéder à la collecte de données dans le cadre d'une EPP peut s'intéresser à l'amélioration de la politique évaluée, son comportement sera aussi dicté par des considérations de marketing stratégique sans rapport avec cette dernière et pouvant l'incliner à édulcorer certains résultats évaluatifs et à en

souligner d'autres (pour fidéliser le client principal par exemple). Il en va de même du côté des commanditaires de l'EPP.

Encadré 2

Illustration : le recours à l'EPP par la « Commission de la production et des échanges »

Entre 1984 et 1985, la « Commission de la production et des échanges » de l'Assemblée Nationale s'est offert les services d'un bureau d'études socio-économiques, filiale de la Caisse des dépôts et consignations, pour évaluer la politique gouvernementale de fermeture définitive des mines de charbon en France.

Une première hypothèse pour expliquer ce phénomène réside dans le fait que Gustave Ansart, le président de ladite commission et député communiste du Nord avait intérêt à répondre favorablement à la demande de l'intersyndicale des mineurs de Charbonnages de France concernant la mise en place d'une mission d'information sur l'intérêt de cesser d'extraire du charbon dans notre pays. Cependant, alors qu'en janvier 1985, M. Ansart, à l'instar des quatre ministres communistes du gouvernement Fabius, démissionne, le projet de mission d'information non seulement n'est pas abandonné, mais devient même un véritable projet d'évaluation avec cahier des charges, appel d'offres et sélection d'un prestataire d'évaluation.

L'enquête (Spencehauer 1995) montre en fait qu'il existe à ce moment-là un petit groupe de jeunes hauts fonctionnaires de l'Assemblée nationale, qui s'inspirant du *Government Accountability Office* du Congrès américain souhaitent doter l'Assemblée d'une certaine autonomie d'expertise et de jugement par rapport au gouvernement et plus largement à la branche exécutive de l'Etat, d'où la mise en scène d'un recours officiel à de l'expertise non-étatique. L'Office Parlementaire d'Evaluation des Choix Scientifiques et Technologiques, créé en 1983, constitue le principal fruit de cette mouvance autonomiste, et l'évaluation de la politique charbonnière de 1984-85 représente un autre de ses coups d'éclats. Bien entendu, cette évaluation n'a eu aucune incidence sur son objet, et pour cause ce n'était ni son but principal, ni son utilité espérée.

Il est également nécessaire de prendre en compte la diversité des motivations des chercheurs participant à des démarches d'évaluation. Le

sociologue James Coleman peut être considéré comme le parangon sincère et intransigeant de sciences sociales axiologiquement neutres mises au service de la décision publique juste via la pratique de l'EPP (Coleman 1966, 1985). Mais comment ne pas établir de lien entre l'investissement de Milton Friedman dans l'évaluation des *negative income tax programs* ou des *school vouchers* (Lacheret 2013) et ses sympathies anti-étatistes et libertariennes ? Aux yeux de certains chercheurs ayant pris part à l'expérimentation du RSA le gauchissement scientifique de celle-ci a été largement compensé par des facilités d'accès à des terrains et des jeux de données jusqu'alors verrouillés et, un peu plus tard, par une participation à l'aventure du Fonds d'expérimentation pour la jeunesse (FEJ), toutes choses qui, d'un strict point de vue scientifique, sont très appréciables.

Les ressorts poussant une minorité non-négligeable de chercheurs en sciences sociales à s'engager dans la pratique de l'EPP ne relèvent que partiellement d'une volonté de bonifier la gestion des affaires publiques au moyen de leur science. Ce décalage entre un discours de résignation devenu convenu et une pratique faisant en réalité peu de cas de la fréquente inutilité gouvernementale de l'EPP est bien assumé par les chercheurs-évaluateurs d'aujourd'hui. Après des décennies de professionnalisations parallèles du savant et du politique, il leur est facile d'arguer que la responsabilité de la faible valorisation gouvernementale de la plupart des EPP incombe essentiellement aux acteurs politico-administratifs impliqués.

Conclusion et ouvertures

L'EPP est née et s'est développée rapidement aux Etats-Unis dans les années 1960, engendrant rapidement frustration et déception. A quelques rares exceptions près (ex. le Coleman Report de 1966), les prescriptions des chercheurs-évaluateurs ne semblent guère suivies par les décideurs politiques. Cette frustration donne naissance aux *Utilization Studies*, menée notamment par Carol H. Weiss. Celle-ci conseille modestie et patience,

engageant avant tout les évaluateurs à produire des résultats satisfaisants d'un point de vue scientifique. Plus entreprenant, et surtout très sceptique sur les effets réformateurs de long terme des évaluations récapitulatives classiques (Wesley et al. 2006), Michael Q. Patton recommande aux chercheurs de descendre dans l'arène politique du système d'acteurs sous-jacent à la politique à évaluer. Et ce, afin d'interagir avec ses tenants et de rassembler les éléments nécessaires à la construction d'un appareil adapté d'évaluation récapitulative et formative. Le principal problème subsistant avec ce modèle participatif réside dans le fait que les acteurs non-scientifiques montrent souvent une malicieuse tendance à pervertir les chercheurs-évaluateurs lors de la construction de l'appareil évaluatif. En définitive, qu'ils soient de posture weissienne ou pattonienne, les chercheurs-évaluateurs sont incités à chercher dans leurs engagements évaluatifs des bénéfices secondaires autres que le maigre ou improbable bénéfice primaire consistant à susciter, grâce à leur science, du changement positif dans l'action publique. Par principe d'autoréalisation, ce comportement rationnel induit une baisse globale de l'utilité gouvernementale réelle des EPP.

Comment sortir de cette situation ? Une solution pourrait résider dans l'assouplissement de la stricte séparation du savant et du politique dans la pratique de l'EPP et de son développement. Car l'expérience montre qu'en matière d'EPP, il n'est guère possible de produire des travaux scientifiques à portée politique réelle sans faire (un peu) de politique. Il ne s'agit pas de manquer de respect aux règles de l'Etat de droit, mais plutôt d'accepter en toute conscience professionnelle de participer à des jeux de pouvoirs, avec les ressources à disposition, afin que moins d'évaluations ne soient condamnées à l'inutilité pour cause de défauts technoscientifiques de fabrication. Les chercheurs, en situation d'EPP, devraient accepter de se prêter davantage au jeu de la technocratie, c'est à dire de consacrer du temps à expliquer, à persuader, à construire des coalitions portant des exigences technoscientifiques sine qua non, saisissant dans l'actualité des

opportunités pour faire avancer des positions réformistes astucieuses... plutôt que de demeurer dans la position confortable selon laquelle « le scientifique propose, le politique dispose ».

Encadré 3

Illustration : les ressources sine qua non à une évaluation utile du « Crédit d'impôt pour la transition énergétique » (CITE, ancien CIDD)

Une évaluation utile du CITE suppose notamment :

- de procéder à une enquête en face-à-face auprès d'un échantillon bien apparié de ménages bénéficiaires (et non bénéficiaires) portant notamment sur leur consommation énergétique (factures d'électricité, de gaz ou de fioul à l'appui) ;
- de mettre en place un observatoire des prix des fournitures thermiques concernées et de la main-d'œuvre d'installation ;
- de procéder à une analyse fine de la structure complexe et évolutive des objectifs du CITE et des acteurs concernés ;
- de cartographier les mesures analogues au CITE prises par les collectivités territoriales librement administrées.

Depuis 2005, les pouvoirs publics nationaux injectent entre un et trois milliards d'euros par an dans le CITE (qui profite essentiellement aux trois premiers déciles de revenus¹⁰), sans qu'aucune des prétendues évaluations officielles produites à son endroit n'aient répondu à ces exigences évaluatives listées ci-dessus.

Ce changement de posture individuelle gagnerait en efficacité et en légitimité si la corporation des enseignants-chercheurs en sciences sociales y donnait force et appui. En termes de reconnaissance d'abord, puisqu'actuellement, beaucoup d'EPST (Etablissements Publics à caractère Scientifique et Technologique) déconsidèrent l'investissement évaluatif de leurs chercheurs, à mesure que la course à la performance académique

¹⁰ Donnée qui permet au moins d'établir que le CIDD ne mérite pas son appellation puisqu'à son trépied de développement durable manque au moins un pied, le social. Pire, il est d'intensité négative, antiredistributif.

s'intensifie. En termes de régulation, peut-être faut-il borner en durée les investissements évaluatifs pour mieux les reconnaître.

Mais il est également possible d'imaginer une solution plus ambitieuse, par exemple la création d'une organisation trans-établissements, transdisciplinaire et trans-sectorielle qui serait en charge de la construction et de l'actualisation régulière de positionnements globaux de la corporation des enseignants-chercheurs en matière de pratique de l'EPP. Cette organisation remplirait des fonctions diverses : portail d'accès vers les viviers académiques de chercheurs-évaluateurs, établissement d'un référentiel commun d'EPP, organisation de formations courtes au cadrage et à la commande d'EPP à l'intention des administrations publiques française (Preskill & Torres 1999), réponse à des saisines ministérielles pour examiner des projets d'EPP, auto-saisine sur le principe de la « *sunshine regulation* » (McCraw 1984 ; Crew & Kleindorfer 2002) concernant de prétendues EPP déjà commises par exemple dans le cadre du CCEPP de l'Assemblée nationale, rédaction de « lettres ouvertes » par exemple pour demander que France Stratégie devienne enfin un vrai « hub » évaluatif national comme envisagé avec pertinence dans le rapport Moreau (2012).

Références

- Barbier J.-C. 2012, Evaluations « expérimentales » : quelques leçons de l'histoire, *Politiques Sociales et Familiales*, vol.110, pp.19-31.
- Barbier J.-C. 2014, Les sources intellectuelles de l'évaluation des politiques publiques, *Politique et Management Public*, 31/1 Janvier-Mars, pp.13-30.
- Birnbaum, R. 2000, Policy scholars are from Venus, Policy makers are from Mars, *The Journal of the Association for the Study of Higher Education*, 23:2 pp.119-32
- Bouchikhi I., Campergue P., Cardinet R., Huré M., Kalaydjian C., Moulin F. 2011, *L'impact de l'expérimentation du RSA*, mémoire de Groupe d'analyse d'action publique du Mastère d'action publique de l'Ecole des Ponts ParisTech, pour le compte de l'IGPDE-Minéfi, sld. Spenlehauer V. & Hamelin F., 56p.
- Bozeman B. 1986, The Credibility of Policy Analysis: Between Method and Uses, *Policy Studies Journal*, 14:4, pp.519-39.
- Braybrooke D. & Lindblom C.E. 1963, *A strategy of decision, Policy Evaluation as a Social Process*, New York, The Free Press.
- Campbell D. T. 1969, Reforms as Experiments, *American Psychologist*, 24 (4), 409-429.
- Chanut V., Cousquer Y. (dir.) 2003, *Une évaluation à l'épreuve de son utilité sociale*, Paris, La Documentation française.
- Coleman J.S. et al. 1966, *Equality of Educational Opportunity*, National Center for Educational Statistics, US Department of Health, Education and Welfare, Washington DC, US Government Printing Office.
- Coleman J.S., 1982, The Emergence of Sociology as a Policy Science, in Coser L.A., Larsen O.N., *The Uses of Controversy in Sociology*, New York, the Free Press, pp.253-61.
- Crew M. A & Kleindorfer, P. R 2002, [Regulatory Economics: Twenty Years of Progress?](#), *Journal of Regulatory Economics*, Springer, vol. 21(1), pp.5-22.
- Crouzel M., Jézéquel A., Quémener JM., Terrier P., *Catastrophisme : objectivité et apprentissage institutionnel*, mémoire de Groupe d'analyse d'action publique du Mastère d'action publique de l'Ecole des Ponts ParisTech, pour le compte de du SG-MEEDE, sld. Dedieu F. & Spenlehauer V., 99p.
- Dunn W.N. (ed.) 1998, *The Experimenting Society: Essays in honor of Donald T. Campbell*, New Brunswick (NJ), Transaction Publishers.

- Flyvbjerg B. 2001, *Making social science matter: Why social inquiry fails and how it can succeed again*, Cambridge: Cambridge University Press.
- Fouquet A. 2013, L'évaluation des politiques publiques : état(s) de l'art et controverses, *Revue Française d'Administration Publique*, n°148, pp.835-47.
- Fredericks, K. A., Carman, J. G. & Birkland, T. A. 2002, Program evaluation in a challenging authorizing environment: Intergovernmental and interorganizational factors, *New Directions for Evaluation*, 95(Fall), pp.5-21.
- Greenberg D.H. & Robins Ph.K. 1986, The changing Role of Social Experiments in Policy Analysis, *Journal of Policy Analysis and Management*, 5, pp.340-62.
- General Accounting Office 1987, *Federal Evaluation: Fewer Units, Reduced Resources, Different Studies from 1980*, Washington DC, GAO.
- Guba, E. G. & Lincoln, Y. S. 1989, *Fourth generation evaluation*, Newbury Park, London, New Delhi, Sage.
- Howlett M. & Ramesh M. 1995, *Studying Public Policy: Policy Cycles and Policy Subsystems*, Toronto, Oxford University Press
- Huberman, M. 1989. Predicting conceptual effects in research utilization: Looking with both eyes. *Knowledge in Society*, (International Journal of Knowledge Transfer), 2(3), 6-24.
- Lacheret A. 2013, L'évaluation comme instrument du sens politique : la controverse autour de l'évaluation des schools vouchers américains, *Revue Française d'Administration Publique*, n°148 pp.923-38
- Leviton L.C. & Hugues E.F. 1981, Research on the Utilization of Evaluations: a Review and Synthesis, *Evaluation Review*, 5 (Aug.), pp.525-48
- Lindblom Ch. 1959, The Science of "Muddling-Through", *Public Administration Review*, 19 (Spring) pp.79-83
- Lindblom Ch. 1965, *The Intelligence if Democracy*, New York, the Free Press.
- Lindblom Ch. 1990, *Inquiry and Change, the troubled attempt to understand and shape society*, New Haven (CT), Yale University Press.
- Mandell M.B. 1984, User-Involvement in the Design and Selection of Evaluation Studies: a User-oriented Framework, *Knowledge*, 5 (June), pp. 419-45.
- Matyjasik N., *L'évaluation des politiques publiques dans une France décentralisée : institutions, marché et professionnels*, Thèse pour le doctorat en science politique soutenue à Sciences-Po Bordeaux le 3 décembre 2010 (dir. Andy Smith)

- McCraw T. 1984, *Prophets of Regulation*, Cambridge (Mass.) Harvard University Press.
- Moreau Y et al. 2012, *Pour un Commissariat général à la stratégie et à la prospective*, Paris, La Documentation Française, 85 pages
- Monnier E., 1987, *L'évaluation de l'action des pouvoirs publics*, Economica, Paris.
- Mouterde F. 2009, Audition de François Mouterde, in Assemblée Nationale, *Compte-rendu de la séance n°2 du Comité de contrôle et d'évaluation des politiques publiques*, 16 juillet, pp. 2-10.
- Nutley S., Walter I., Davies H. 2007, *Using Evidence: How Research Can Inform Public Services*, Bristol (UK), the Policy Press (University of Bristol)
- Patton M.Q. 1975, *In Search of Impact: An Analysis of the Utilization of Federal Health Evaluation Research*, Center for Social Research, University of Minnesota, Minneapolis.
- Patton, M. Q. 1982, *Practical Evaluation*, Beverly Hills, London, New Delhi, Sage
- Patton M.Q. 1988, The Evaluator's Responsibility for Utilization, *Evaluation Practice*, 9:2 pp.5-24
- Patton, M. Q. 1997 (1978), *Utilization-focused evaluation*, (3rd ed.), Thousand Oaks, London, New Delhi, Sage
- Pawson R. 2006, *Evidence-Based Policy. A Realist Perspective*, London, Sage
- Preskill, H. & Torres, R. T. 1999, Building capacity for organizational learning through evaluative inquiry, *Evaluation*, 5(1), pp.42-60
- Sabatier P.A. 1978, The Acquisition and Utilization of Technical Information by Administrative Agencies, *Administrative Science Quarterly*, 23 (Sept.) pp.396-417
- Sabatier, P.A. 1988, An advocacy coalition framework of policy change and the role of policy-oriented learning therein, *Policy Sciences*, 2, pp.129-68
- Scriven M. 1967, The methodology of evaluation, in R. W. Tyler, R. M. Gagné, & M. Scriven (Eds.), *Perspectives of curriculum evaluation*, 39-83. Chicago (Ill.), Rand McNally
- Simon H. 1947, *Administrative Behavior*, New York, the Free Press.
- Simon H. 1997, Managing in an Information-Rich World, in Simon H. (ed.), *Models of Bounded Rationality*, Cambridge (Mass.), MIT Press
- Spencehauer V. 1995, *L'évaluation de politique, usages sociaux. Trois études de cas d'évaluation*, Paris, L'Harmattan.

- Spencehauer V. 2013, L'ancrage de l'évaluation des politiques publiques dans les sciences sociales aux Etats-Unis : reconstruction d'un linéament intellectuel, *Revue Française d'Administration Publique*, n°148 pp. 877-93
- Trochim, W.M.K. 2009, Evaluation policy and evaluation practice, in Trochim W.M.K, Mark M. & Cooksy L. J. (Eds.), *Evaluation policy and evaluation practice, New Directions for Evaluation*, 123, 13–32
- Vedung E. 2000, *Public Policy and program Evaluation*, London, Transaction Publishers
- Vedung E. 2010, Four Waves of Evaluation Diffusion, *Evaluation*, 16:3 pp.263–277
- Weiss C.H. 1967 (April), Utilization of evaluation: Toward comparative study, in *House of Representatives committee on government operations, the use of social research in federal domestic programs*, Part III Washington, D.C.: Government Printing Office pp.426-32.
- Warin Ph. 2014, Une recherche scientifique dans le *problem-solving* : un retour d'expérience, *Politique et Management Public*, 31/1 janvier-mars pp.113-122
- Weaver K. 1986, The Politics of Blame Avoidance, *Journal of Public Policy*, 6:4 pp.371-98
- Weiss C.H. 1979, The Many Meanings of Research Utilization, *Public Administration Review*, 39 (Sept-Oct) pp.426-31
- Weiss C.H. 1983, Ideology, Interest, and Information: The Basis of Policy Positions, in Callahan D. and Jennings B. (eds.), *Ethics, Social Science, and Policy Analysis*, New York, Plenum Press, pp. 213-45
- Weiss, C. H. 1990, If program decisions hinged only on information: a response to Michael Patton, in Alkin M. (ed.), *Debates on evaluation* (pp. 208-229), Newbury Park, London, New Delhi, Sage
- Weiss C.H. 1993a, Where Politics and Research Evaluation Meet, *Evaluation Practice*, 14:1, pp.93-106
- Weiss C.H. 1993b, Mega-Lessons from U.S. Evaluation Experience, *Politica*, Bind 25, 12p.
- Wesley F., Zimmerman B., Patton M.Q. 2006, *Getting to Maybe: How the World Has Changed*, Canada, Random House
- Wildavsky A. 1969, Rescuing Policy Analysis from PPBS, *Public Administration Review*, 29 (2) pp.189–202
- Williams W. 1971, *Social Policy Research and Analysis, the Experience of in the Federal Social Agencies*, New York, Elsevier.

2016/03



**Le LIEPP (Laboratoire interdisciplinaire d'évaluation des politiques publiques) est un laboratoire d'excellence (Labex).
Ce projet est distingué par le jury scientifique international désigné par l'Agence nationale de la recherche (ANR).
Il est financé dans le cadre des investissements d'avenir.**

(ANR-11-LABX-0091, ANR-11-IDEX-0005-02)

www.sciencespo.fr/liepp

Directeurs de publication:

Bruno Palier et Etienne Wasmer

Sciences Po - LIEPP
27 rue Saint Guillaume
75007 Paris - France
+33(0)1.45.49.83.61
liepp@sciencespo.fr

