



HAL
open science

Quelle prise en compte des informations pour quelle hiérarchisation des mondes possibles? Une typologie des démarches d'évaluation

Guillaume Hollard, Antoine Vion

► To cite this version:

Guillaume Hollard, Antoine Vion. Quelle prise en compte des informations pour quelle hiérarchisation des mondes possibles? Une typologie des démarches d'évaluation. Cahiers du CEVIPOF, 2006, 44, pp.87-123. halshs-00120502

HAL Id: halshs-00120502

<https://shs.hal.science/halshs-00120502v1>

Submitted on 15 Dec 2006

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.

QUELLE PRISE EN COMPTE DES INFORMATIONS POUR QUELLE HIERARCHISATION DES MONDES POSSIBLES ? UNE TYPOLOGIE DES DEMARCHES D'EVALUATION¹

GUILLAUME HOLLARD

Maître de conférences en sciences économiques à l'Université de Marne-la-Vallée, OEP (en délégation au CNRS au sein du laboratoire EUREQUA)

Et

ANTOINE VION

Maître de conférences en science politique à l'Université de la Méditerranée, LEST (UMR 6123)

¹ Une version préliminaire de ce texte a été présentée au 8ème congrès de l'AFSP à Lyon dans le cadre de l'atelier 20 « L'analyse des politiques publiques entre science des modèles et sciences de l'enquête ». Nous remercions les organisateurs de cet atelier (Pascale Laborier et Stéphane Nahrath), ainsi que les discutants (Jean-Claude Passeron et Jean-Louis Fabiani) et les autres participants. Pierre Lascoumes a bien voulu nous en faire un commentaire détaillé, les améliorations apportées lui doivent beaucoup. Nous remercions aussi Philippe Mossé, Francesca Petrella, et les participants au séminaire d'économie des conventions (Olivier Favereau, Franck Bessis, François Eymard-Duvernet, parmi d'autres) qui ont bien voulu se livrer à des discussions détaillées des versions successives.

La pratique de l'évaluation des politiques publiques tend à se diffuser sous les formes les plus diverses. Il suffit, pour s'en rendre compte, de comparer les revues spécialisées américaines, telles que l'*American Journal of Evaluation* ou *Practical evaluation* ou encore les publications gouvernementales et intergouvernementales telles que les revues de l'OCDE ou de la Banque Mondiale. Pour la France, les lectures d'*Economie et statistiques* ou d'*INSEE Premières* peuvent être utilement croisées avec celles de revues telles que *Politiques et Management Public* ou *Economie Publique*. Les différents développements de l'évaluation sont désormais bien connus (Duran, Monnier, 1992 ; Kessler, Lascoumes, Setbon, Thoenig, 1998 ; Spenlehauer, 1999 ; Varone & Jacob, 2003, 2004a, 2004b, etc.). Les controverses récurrentes, mais rarement étalées au grand jour, qui traversent ce champ de l'évaluation des politiques publiques, peuvent être vues comme une illustration supplémentaire des incompréhensions réciproques entre les tenants de la modélisation économique ou économétrique et les praticiens de l'enquête qualitative. Comme le soulignait récemment Patrice Duran, « l'évaluation a moins à craindre de son institutionnalisation proprement dite que de son enfermement dans des discours d'école » (Duran, 2002).

Notre propos n'est pas d'historiciser l'analyse des controverses entre écoles ou des pratiques des professionnels (Fourquet, 1976 ; Spenlehauer, 1999, 2003 ; Desrosières 2003a), ni de jeter un regard critique sur les postures des évaluateurs (Amiot, 1986 ; Mény, Thoenig, 1989 ; Fontaine, 1996 ; Spenlehauer, 1995, 1998, 2000 ; Thoenig, 2002). Nous défendons plutôt la thèse selon laquelle certaines controverses peuvent être mieux comprises si l'on explicite les manières de traiter l'information utilisée pour la démonstration.

Selon notre thèse, le traitement de l'information opéré dans le but de produire une expertise, et plus généralement un raisonnement en sciences sociales, peut être séparé en deux opérations logiques élémentaires : la prise en compte et la hiérarchisation. Nous nous inspirons sur cette question de la distinction établie par Latour (1999) entre prise en compte et ordonnancement à propos des modalités de la décision publique, mais nous proposons des définitions un peu différentes². Nous entendons ici par prise en compte la manière d'inclure des informations nouvelles dans le raisonnement. L'opération de prise en compte porte donc sur des informations. La hiérarchisation est une opération consistant à utiliser les informations prises en compte pour produire une hiérarchie des mondes possibles.

Constatons que ces opérations sont banales et communes à l'ensemble des raisonnements mobilisés dans le cadre d'expertise. Un antagonisme entre ces deux opérations banales apparaît lorsqu'il convient d'explicitier le raisonnement suivi. Certains styles d'évaluation n'auront pas de mal à donner des gages de transparence sur les opérations de prise en compte mais laisseront dans l'ombre la manière dont le travail de hiérarchisation a été effectué. Réciproquement, d'autres styles d'évaluation seront à même d'explicitier complètement leur manière de hiérarchiser tandis que les opérations de prise en compte seront beaucoup moins vérifiables par un tiers. Avoir recours à un style d'évaluation impose un choix sur ce qui pourra être explicité. Les antagonismes et autres controverses qui traversent le champ de l'expertise pourraient avoir leur source dans une opposition logique qui tient à l'impossibilité de produire des raisonnements totalement transparents qui épouserait les canons des sciences non humaines.

Nous allons exemplifier notre thèse en analysant en détail quatre grands styles d'évaluation, définis dans le tableau suivant : la démarche utilitariste (1), la démarche instrumentale (2), la démarche praxéologique (3) et la démarche procédurale (4). Nous montrerons, pour chacune comment

² Le pouvoir de prise en compte doit permettre de déployer la « pluralité des mondes » grâce à des procédures, des « enquêtes pour détecter, rendre visible, et faire parler les différentes propositions candidates à l'existence ». Il suppose la mise en place de procédures de consultation pour constituer « le jury capable de juger des effets de chaque proposition sur les habitudes des autres ». Le pouvoir d'ordonnancement répond quant à lui au besoin d'unification, d'institution d'un monde commun. Il suppose une exigence de clôture (au moins temporaire) des controverses et des procédures contradictoires pour composer la hiérarchie optimale, c'est-à-dire finalement, décider.

procède le raisonnement afin de mettre en évidence la manière dont s'opèrent les opérations de prise en compte et de hiérarchisation.

La suite du document passe successivement en revue chacune de ces démarches, d'abord à un niveau général, puis en exemplifiant le raisonnement. En conclusion, nous étayerons notre thèse au regard des éléments rassemblés.

	UTILITARISTE	INSTRUMENTAL	PRAXEOLOGIQUE	PROCEDURAL
PRINCIPE	Analyse des préférences individuelles et des usages de biens	Réorganisation par les outils de gestion	Compréhension des pratiques administratives et des pratiques d'usagers	Irréductibilité des valeurs et primat de la délibération pour désamorcer les conflits et parvenir à un accord
STYLE DE DEMARCHE EVALUATIVE	1. CONSTITUTION D'UNE BASE DE DONNEES OU D'UNE LISTE DE VARIABLE 2. CONSTRUCTION DE L'ENJEU ET OBJECTIVATION DE LA SITUATION SUR LA BASE DES DONNEES 3. MISE EN CONCURRENCE D'EXPLICATIONS CAUSALES PAR TESTS DE VARIABLES 4. HIERARCHISATION DES MONDES POSSIBLES 5. DEFINITION D'UN OPTIMUM	1. CLASSIFICATION D'ACTIVITES INDIVIDUELLES OU COLLECTIVES (NOMENCLATURE) 2. VALORISATION DES ACTIVITES PAR INDEXATION 3. COMPARAISON DES VALEURS AGREGÉES 4. HIERARCHISATION DES RESULTATS	1. DEFINITION D'UN ENSEMBLE DE PRATIQUES A ANALYSER 2. RECUEIL DE DONNEES SUR LA BASE D'HYPOTHESES 3. COMPARAISON DES EXPLICATIONS CAUSALES ISSUES DE L'ENQUETE 4. HIERARCHISATION DES EXPLICATIONS SOUS LA FORME D'UN CADRAGE THEORIQUE 5. DIAGNOSTIC DES CAUSES DE SUCCES OU D'ECHEC	1. DEFINITION D'UN ENJEU OU D'UN PROBLEME 2. MISE EN PLACE D'UNE PROCEDURE DE CONSULTATION 3. PRISE EN COMPTE DE L'ENSEMBLE DES AVIS ARGUMENTES 4. SELECTION DES BONS ARGUMENTS 5. SYNTHÈSE DES BONS ARGUMENTS
TYPES D'ENQUETES CONSTRUITES	- Recueil de données par questionnaires - Le cas échéant, enquêtes comportementales expérimentales ou en situation	- Collectes de données comptables - Création d'indicateurs - Recherche, instruction et sanction des fautes de gestion	- Enquêtes de fréquentation (statistiques descriptives) - Echelles d'attitudes - Enquêtes par entretiens - Observations directes - Recherche-action	- Collecte des arguments contradictoires - Recherche de preuves
ENJEUX DE DEBAT	- Evolution des préférences - Représentativité des enquêtés - Existence d'une variable manquante	- Pertinence des indicateurs - Imputation comptable et imputabilité politique des dépenses publiques	- Epreuve de réalité des faits exemplifiés - Interprétation des observations	- Légitimité des dispositifs, des porte-parole, des arguments
PLACE DU POLITIQUE	Mise en œuvre des recommandations issues de l'expertise	Arbitrage budgétaire entre les catégories pré-définies	Interprétation et traduction dans le langage commun	Possibilité d'ignorer les résultats de l'expertise ; rivalité dans la mise en forme de la volonté populaire
PRISE EN COMPTE D'INFORMATIONS OU D'ACTEURS NOUVEAUX	Faible voire nulle sans modification de la procédure	Limitée aux catégories existantes	bonne	Très bonne car pas de contrainte a priori
DEGRE DE HIERARCHISATION	FORT	MOYEN-FORT	MOYEN	Faible

1. LA DEMARCHE UTILITARISTE

Lorsque nous employons ici l'adjectif « utilitariste », nous ne renvoyons pas seulement au courant de pensée utilitariste suivant une stricte approche historique des idées économiques. Nous envisageons plutôt ce terme comme le qualificatif d'une démarche générique centrée sur les comportements individuels (par opposition à l'observation de situations d'interaction). Les opérations d'évaluation qui peuvent être menées dans cette logique consistent à construire des

modèles mathématiques, et à les soumettre à une série d'épreuves. Nous nous concentrerons sur la nature particulière du couplage opéré entre hiérarchisation et prise en compte dans ce type de démarche. Nous plaiderons l'existence d'une séquentialité qui initie le raisonnement en recensant toutes les informations à prendre en compte avant d'élaborer un raisonnement hiérarchisant ces informations. Nous commençons par décrire le raisonnement en termes généraux avant de l'exemplifier par un article récent d'analyse des politiques publiques, en l'occurrence un article de Thomas Piketty portant sur l'influence de la taille des classes sur la réussite scolaire.

Dans ce type de démarche, les éléments pris en compte sont limités et exprimés dans des catégories prédéfinies au sens où un sondage en question fermée contraint au préalable la nature de l'information recueillie. Les éléments se présentent à l'issue de ce travail sous forme d'une base de données ou d'une liste de variables. Une des caractéristiques de ce mode de raisonnement est de ne discuter l'adéquation entre la chose étudiée et sa mesure qu'à titre de préliminaire. L'analyse proprement dite est distincte de cette phase de discussion. Le travail de prise en compte, ainsi que la discussion critique inhérente à celui-ci, doivent être clos avant de passer à l'analyse proprement dite. Les informations sont listées mais une seconde phase, moins visible, consiste à réaliser un appariement entre les données ou variables utilisées et les notions manipulées. Par exemple une analyse comparée des systèmes productifs de différents pays s'appuie sur la comparaison de données particulières : les niveaux de PIB. Le calcul du PIB lui-même se fonde sur une longue série de traductions et classifications des transactions marchandes. Les insuffisances du PIB sont largement soulignées et abondamment commentées : le débat occupe une large part des programmes de première année d'économie et figure en bonne place dans les manuels élémentaires de cette discipline. A titre d'exemple, on souligne souvent que les transactions qui franchissent la frontière entre le marchand et le non-marchand altèrent la mesure du PIB. Si monsieur épouse sa femme de ménage (ou madame son chauffeur), le PIB diminue. Que dire par ailleurs d'un pays où le troc est largement répandu et la production agricole largement autoconsommée ? On conçoit bien que résumer en un seul chiffre l'ensemble du système économique national suppose quelques approximations. Quand il s'agit ensuite d'utiliser les données ou les variables retenues, la discussion est close. Par convention, les termes « PIB » et « richesse produite » sont devenus équivalents. On prête à Alfred Binet, l'inventeur du premier test de QI cette phrase emblématique : « L'intelligence ? C'est ce que mesure mon test ». On pourrait dire de manière similaire : « La richesse ? C'est ce que mesure le PIB ». On instaure donc une correspondance biunivoque entre une notion abstraite (l'intelligence, la richesse) et sa mesure (le QI, le PIB). La phase de prise en compte légitime, ou du moins tente de rendre légitime, un usage des données comme si celles-ci suffisaient à capturer l'essentiel du contexte duquel elles sont extraites.

Une fois close l'opération de prise en compte, il est possible de définir l'état du monde présent, sous la forme d'écart de richesse entre les pays, d'inégalités des niveaux d'intelligence, etc. Le travail du modélisateur consiste alors à élaborer un modèle explicatif qui permet de définir les modalités d'intervention sur ce monde présent pour produire un éventail de mondes possibles. Ces mondes possibles sont ensuite hiérarchisés au regard de critères explicites (critère rawlsien, critère égalitariste, critère d'efficacité, etc.). Le modélisateur se place dans une position de neutralité par rapport aux valeurs puisque l'opération de hiérarchisation se fait au nom de critères explicites et extérieurs au modélisateur. C'est la partie noble de son raisonnement, là où s'exprime pleinement le talent de celui qui conduit l'expertise. Le chemin qui conduit de la prise en compte initiale aux conclusions du raisonnement est clairement balisé et, idéalement, ne laisse rien dans l'ombre. Il est possible de produire un raisonnement transparent au sens où celui-ci peut-être conduit au nom de la logique seule. Dans cette seconde étape, le raisonnement épouse les canons de la science : il est logique, souvent formel et peut être reproduit à l'identique par un tiers. Il donne donc des garanties de transparence qui, associées au caractère quantitatif de l'analyse, rendent l'expertise difficilement discutable.

Ce style d'évaluation produit des résultats de la forme « si X alors Y ». Typiquement, « si les objectifs de la décision sont ceux-ci (réduire le chômage, diminuer les inégalités, etc) alors il convient de mettre en œuvre telle politique (mettre en œuvre un nouveau type de contrat de travail, accroître tel

type de financement, etc.». Les destinataires de l'évaluation n'ont d'autre choix que d'appliquer, ou non, les recommandations résultant de l'évaluation. Charge à eux de les rendre légitimes, ou au moins acceptables.

Venons en à notre exemple, concernant l'impact des tailles de classes sur la réussite scolaire. Thomas Piketty présente dans cette étude³ les résultats d'une analyse économétrique à partir d'une base de données collectée par la Direction des Etudes et de la Prospective du ministère de l'Education Nationale. Nous avons retenu cet article pour exemplifier notre propos général du fait qu'il s'agit d'un travail largement salué comme ayant renouvelé la vision de ce sujet classique. Il est considéré comme une bonne évaluation. L'annonce du plan résume les grands traits de la démarche :

« La suite de cette étude est organisée de la façon suivante. La section 2 décrit les principales caractéristiques techniques des données utilisées. La section 3 plante le décor en présentant un certain nombre de statistiques descriptives concernant la taille des classes et les inégalités de réussite scolaire dans les écoles primaires françaises. La section 4 présente les résultats de nos estimations de l'impact de la taille de la classe sur la réussite scolaire, en commençant par les estimations « naïves », puis en développant la méthode fondée sur les discontinuités liées au franchissement des seuils d'ouverture des classes. Nous utilisons ensuite les résultats de ces estimations dans la section 5 pour évaluer dans quelle mesure des politiques de ciblage en faveur des écoles défavorisées plus ambitieuses que celles menées aujourd'hui pourraient permettre de réduire les inégalités scolaires. Dans la section 6, nous présentons l'impact de la ségrégation sociale sur la réussite scolaire, afin notamment de comparer l'ampleur de ces effets à ceux des tailles de classes réduites. Enfin, la section 7 présente un certain nombre d'extensions et de questions ouvertes. »

La partie 2 de l'article décrit les informations contenues dans la base de données et la partie 3 ajoute d'autres informations à prendre en compte. Il est important de noter que l'apport d'informations nouvelles est limité à ces deux sections du texte. L'appariement entre notions (la réussite scolaire, les élèves défavorisés, etc.) et mesures est fixé à ce niveau. Par la suite la réussite scolaire sera assimilée ce que mesurent les tests de la DEP, la notion d'élève défavorisé se confondra avec la PCS des parents, etc. De plus, les marqueurs du contexte disparaissent. On ne parle plus d'un échantillon d'école à une date donnée mais du système scolaire français en général.

Ces opérations d'appariements et de décontextualisation sont caractéristiques des opérations de prise en compte dans ce style de démarche. Le point important pour notre propos est de noter que ces opérations recèlent une part d'ombre qui, quelque soit la bonne foi de l'évaluateur, ne peut être rendue complètement transparente. On s'en remet nécessairement à la fiabilité supposée des institutions qui ont produit les données, au prestige et à la renommée de l'auteur.

Une fois ce préalable franchi, le raisonnement devient transparent (l'auteur rend même les données accessibles sur son site Internet). Chacun peut donc combiner les données comme il lui convient. L'auteur utilise un modèle économétrique simple. L'apport de l'analyse est d'utiliser l'effet de seuil introduit par l'existence d'un seuil d'ouverture ou de fermeture de nouvelles classes. Ainsi, il peut observer dans un environnement social constant, l'effet d'une réduction ou d'une hausse du nombre d'élèves par classe. Reproduire les mêmes calculs que l'auteur conduira aux mêmes résultats. La section 5 est emblématique de la démarche décrite : le modèle est étendu à des situations hypothétiques (ce qui se passerait si l'on réduisait la taille des classes). Ces situations sont comparées entre elles au regard d'un critère rawlsien qui se concentre sur la situation des plus défavorisés (une situation est d'autant meilleure que la situation des plus défavorisés est meilleure). L'optimum rawlsien est calculé de manière formelle. L'auteur en tire des recommandations pratiques et chiffrées : une réduction de 40% des inégalités constatées peut-être atteinte, à moyens constants, en diminuant la taille des classes en ZEP jusqu'à 18 élèves (contre 22 actuellement) au prix d'une légère remontée de la taille des classes hors ZEP, qui passeraient de 23.3 à 24.2 élèves. Le message final est bien de la forme « si vous souhaitez réduire les inégalités scolaires, alors voici ce

³ Nous nous basons ici sur un document de travail accessible en ligne : <http://www.jourdan.ens.fr/piketty/fichiers/public/Piketty2004b.pdf>

qu'il convient de faire ». On peut difficilement être plus prescriptif. La section 6 expose un modèle théorique prenant en compte simultanément deux sources d'inégalités liées à la taille des classes et la ségrégation scolaire. La section 7 boucle le raisonnement en pointant les informations manquantes qui permettraient d'étendre la conclusion.

Les critiques adressées à style d'expertise sont connues, et presque imparables. Au regard de ce qui précède, il est possible d'anticiper le fait qu'elles portent principalement sur la phase de prise en compte puisque que c'est à ce niveau qu'interviennent les (inévitables) bricolages méthodologiques. L'exemple le plus courant concerne l'existence d'une causalité non prise en compte. Imaginons par exemple que l'enseignement en ZEP soit plus axé sur une pédagogie participative qui vise le développement harmonieux des enfants plutôt que leur capacité à réussir des tests formels. L'outil de mesure de la réussite scolaire retenu (des tests standardisés) trouverait donc sans surprise un meilleur résultat hors ZEP et ne ferait en fait que constater un écart de méthode pédagogique, et non pas un écart de réussite scolaire. Dans cet exemple (fictif), l'adéquation entre les données exploitées et la notion de réussite scolaire peut également être mise en doute puisque la réussite scolaire peut ne pas prendre le même sens en ZEP et hors ZEP. La méthode d'expertise employée ne permet pas de se prémunir contre des critiques de ce type, alors qu'il est possible de reproduire les calculs qui ont mené aux conclusions.

Le principe d'une évaluation portée par un raisonnement quantitatif a une longue histoire en France et remonte au moins aux ingénieurs économistes tels que Cournot ou Dupuit. Cette tradition de serviteur de l'Etat soucieux de développer les outils d'une véritable science économique s'est prolongée dans la création en 1930 de l'influente société d'économétrie. Il en a résulté un recours privilégié au raisonnement hypothético-déductif qui a durablement marqué le champ de l'évaluation. L'analyse *ex-post*, qui inclut l'exemple traité ci-dessus, trouve sa pleine justification dans l'analyse d'expérience naturelle. On regroupe sous ce terme des situations particulières où une partie de l'échantillon a été soumise à une transformation sociale et l'autre non. Par exemple, on peut comparer deux cohortes de chômeurs qui ont été aléatoirement affectées à deux dispositifs d'aide à l'emploi différents (Cahuc, Zylberberg 2004) ou comparer les bénéficiaires et les non bénéficiaires de dispositifs de formations de chômeurs pour en mesurer les impacts (Van der Linden, 2001). De ce fait, les variables manquantes peuvent être supposées égales d'un groupe à l'autre, ce qui rend l'opération de prise en compte moins vulnérable. Les expériences naturelles fournissent un point d'appui privilégié puisque le raisonnement devient difficile à remettre en cause. La place laissée au politique est par conséquent restreinte à la mise en œuvre des recommandations de l'expertise.

Cependant, l'analyse peut également opérer *ex-ante* et l'arsenal mis en place n'est pas moins impressionnant dans ce cas. Le calcul économique, abondamment employé en matière de transport, d'environnement et, plus récemment, de santé, a élaboré en effet un ensemble de dispositifs dont l'objectif affiché est de pouvoir rendre comparables *ex-ante* les différentes dimensions d'une décision publique, pour pouvoir *in fine* hiérarchiser les différents projets. La hiérarchisation opère en rabattant chacune des dimensions de la décision sur un axe unique, gradué en unité monétaire. Un haut degré de sophistication est atteint pour fournir des équivalents monétaires à des aspects aussi peu monétisables a priori que la qualité de vie (Geniaux et al. 2001) ou les effets locaux de la pollution atmosphérique (Nicolas et al, 2002). Les développements récents vont jusqu'à enquêter sur les préférences futures des agents pour donner une valeur à des politiques à mettre en œuvre. L'évaluation contingente est emblématique de cette approche. Elle marque une évolution vers la prise en compte de données issues d'interviews, alors que l'analyse traditionnelle reposait sur la prise en compte d'actions observées ou de quantités objectives. En dépit de ses défauts (Flachaire et Hollard, 2006), la méthode d'évaluation contingente connaît un développement rapide à travers le monde.

S'agissant d'anticiper des pratiques futures, l'évaluation *ex-ante* se base sur des informations moins solidement établies que dans le cas d'une expérience naturelle. De ce fait, la place accordée au politique par les experts diffère quelque peu et semble évoluer vers une reconnaissance du rôle des

élus dans la détermination des choix futurs (Crozet 2004). A titre d'exemple, le « coût du mort » qui permet de déterminer l'opportunité de nouveaux investissements en matière de sécurité, conduisait à disqualifier les choix de certains élus qui semblaient sur-investir dans des équipements de sécurité (« zone 30 », ralentisseurs, etc). Or, quelques années plus tard, le coût du mort ayant été réévalué, ces mêmes équipements étaient devenus rentables au regard des nouveaux critères en vigueur.

2. LA DEMARCHE INSTRUMENTALE

Nous entendons par démarche instrumentale d'évaluation toute entreprise visant à construire un dispositif technique produisant des valeurs sur la base d'outils de gestion⁴ destinés à optimiser la gestion des moyens d'une organisation. Ce type de dispositif repose donc sur la réalisation et la routinisation de procédures automatiques de catégorisation et de comptage qui forment a priori la hiérarchie des informations pertinentes pour conférer sa valeur à l'activité observée. Ce style de démarche évaluative se caractérise donc par le fait que des classifications sont établies a priori sous forme de nomenclatures, chaque item étant ensuite valorisé par indexation (voir encadré pour l'exemple qui suit), puis les valeurs agrégées comparées et les agrégats hiérarchisés.

Trois points de discussion nous semblent importants pour comprendre ce qui est en jeu dans ce type de démarche : le travail de construction des modes de sélection des informations, le caractère difficilement réversible de l'instrumentation, et les effets de standardisation des comportements qu'elle engendre.

Le travail de construction des modes de sélection des informations est toujours, dans ce type de démarche, l'objet de vifs débats entre les techniciens chargés de configurer l'outil, les autorités politiques et administratives commanditaires, et les services dont la gestion sera affectée par l'outil. Pour ne prendre qu'un exemple récent, la mise en œuvre de la LOLF⁵ dans les administrations nationales françaises met au centre des discussions la pertinence des informations susceptibles de servir à des opérations de mesure de la performance⁶. Faut-il, par exemple, pour évaluer l'activité d'une équipe d'enseignants d'une UFR, définir le pourcentage du service individuel effectué en recherche, en enseignement, et en administration, et faire une moyenne par UFR, par statuts, etc. ? Ce mode de mesure pose le double problème de la catégorisation des tâches (ne prépare-t-on pas un cours lorsqu'on cherche et réciproquement ?) et de la mise en équivalence des statuts : doit-on s'en tenir aux grades et échelons ou faut-il introduire des variables plus fines comme le taux d'encadrement par établissement, voire par formation, ce qui nuancerait très fortement le postulat selon lequel la charge d'enseignement d'un maître de conférences enseignant par exemple dans un diplôme de Licence de psychologie équivaldrait à celle d'un maître de conférences en mathématiques. Ces questions ne sont évidemment pas détachées des anticipations que font les protagonistes des usages de l'instrument en termes de mesure des performances⁷.

Si le travail de construction de l'instrument d'évaluation n'est pas neutre, il n'est pas non plus complètement libre, dans la mesure où il est conditionné par la résolution de contraintes techniques parfois difficilement surmontables. Que se passe-t-il, dans notre exemple précédent, si le travail de programmation informatique permet la prise en compte de la catégorie « taux d'encadrement par diplôme », mais que le travail de modélisation économétrique destiné à la comparaison des

⁴ On appellera outil de gestion « un ensemble d'outils formalisés (textes, logiciels...) destinés à coordonner des opérations de gestion (du travail, des produits, des processus, des personnels...) au moyen de procédures considérées comme efficaces » (Boussard, Mercier, Tripier, 2004).

⁵ Pour une présentation technique approfondie de la LOLF, voir le numéro spécial de la Revue française de finances publiques de 1999.

⁶ Pour une discussion de la mise en œuvre de la LOLF dans la magistrature, cf. Pallez, 2002

⁷ Alain Desrosières (1992) montre ainsi, par exemple, que le travail des statisticiens est perpétuellement traversé par la tension entre le souci d'optimiser la mesure et la peur panique de développements potentiels d'usages politiques non maîtrisables⁷. La sociologie des sciences inspirée par Simondon a montré que la construction d'instruments, loin d'être une technicité neutre, met constamment en jeu l'arbitrage entre des voies de développement différentes.

performances ne permet pas de prendre simultanément en compte l'ensemble des taux par diplôme pour mesurer la charge de travail effective d'un enseignant et d'agrèger ensuite ces charges par établissement ? Certains ont à juste titre insisté sur le fait que la sélection des informations opère une réduction de la complexité organisationnelle. Mais il faut aussi souligner, en écho aux débats de sociologie des sciences sur le statut ontologique des instruments, que les limites techniques éprouvées dans la construction du dispositif affectent le sens même de l'activité évaluée. S'il est donc toujours possible de produire des réinterprétations par la production de nouvelles convergences et mises en relation des informations (Akrich, Callon, Latour, 1988), cela ne met qu'exceptionnellement en cause la légitimité de l'instrument lui-même, qui tend à s'imposer progressivement à l'ensemble de la communauté intéressée (Desrosières, 1993 ; Bağla-Gökalp, 1996), à moins d'arbitrages politiques décidant le recours à des formes d'instrumentation alternatives exclues par le passé (Barthe, 2003).

Tout un pan de la recherche sur l'action publique s'interroge aujourd'hui sur ce phénomène de « durcissement » progressif des instruments d'évaluation⁸. L'exemple de l'évaluation de la gestion hospitalière en France est sans doute le plus significatif en la matière. La Direction Générale de la Santé a introduit au début des années quatre-vingt dans la gestion hospitalière un instrument d'analyse des séjours d'hospitalisation destiné à évaluer l'activité des établissements en comparant les types de pathologie et les types de durée de séjour des patients. Cet instrument, le Programme de Médicalisation des Systèmes d'Information (PMSI), était inspiré de travaux américains réalisés le Professeur Fetter et son équipe à l'Université de Yale, consistant à recouper systématiquement des informations sur les malades et leur traitement en *diagnostic related groups* (DRG). Dans cette logique, le PMSI a défini, pour les activités dites « de court séjour » (médecine, chirurgie et obstétrique)⁹ des opérations de recouplement des informations médicales (identification des types de pathologie et de traitement en groupes homogènes de malades (GHM)) et des informations administratives (détermination de la longueur du séjour, calcul du prix de journée) permettant aux établissements hospitaliers de rationaliser la gestion des lits, et de comparer leurs performances. Au moment du lancement officiel de cette réforme, au milieu des années quatre-vingt, le Directeur des Hôpitaux présentait le PMSI, à l'issue de nombreux débats, comme une simple amélioration du système d'information sur les malades permettant un contrôle interne de la gestion des lits. Si un instrument comme le PMSI s'est progressivement imposé, c'est parce qu'il offrait aux médecins des hôpitaux les moyens de se positionner par rapport aux cliniques et de comparer entre eux les effets de leurs techniques de soin sur la gestion des lits. C'est parce que l'instrument d'évaluation est aussi un instrument de valorisation d'une communauté professionnelle qu'il tend à s'imposer. Comme le suggère Giddens (2000), la modernisation n'est pas seulement un processus de rationalisation technique dominant, c'est aussi un processus qui développe la réflexivité des individus sur leurs pratiques.

La démarche d'évaluation a ensuite été précisée en 1991 par une disposition réglementaire obligeant les hôpitaux à constituer pour chaque patient un résumé de sortie standardisé (RSS) recoupant un diagnostic principal, des diagnostics associés, des actes techniques, l'âge, le sexe, etc. L'introduction d'un logiciel groupeur a ainsi permis d'opérer une nomenclature de 540 produits à partir de laquelle les prestations hospitalières devenaient comparables. C'est la construction de cette comparabilité des données qui a ensuite permis aux économistes de la santé, et particulièrement les économètres, de concevoir des modèles d'agrégation permettant l'indexation de l'activité des établissements sur un standard de performance. La construction de ce nouveau savoir a ensuite permis à l'administration des hôpitaux, dans le cours d'une réforme de décentralisation créant les agences régionales d'hospitalisation, d'opérer une valorisation économique de l'activité. Cette valorisation a pris la forme suivante : « la valorisation économique relative des GHM est ensuite construite à

⁸ On peut utilement faire un point sur cette question en se reportant à l'introduction de Pierre Lascoumes et Patrick Le Galès, et aux contributions de Philippe Bezès et Patrick Le Galès dans le livre *Gouverner par les instruments*, paru l'an dernier aux Presses de Sciences PO (Lascoumes, Le Galès, 2004).

⁹ La tendance actuelle est à l'extension des domaines d'activité couverts par le PMSI.

partir d'un échantillon d'établissements participant à ce qu'il est convenu d'appeler « l'étude nationale des coûts ». Ces établissements publics ou participant au service public disposent tous (Charretton et Fischesser, 1995) d'une comptabilité analytique détaillée rendant possible le calcul du coût complet de chaque séjour¹⁰. Lorsque la base des coûts des séjours est constituée, il ne reste plus qu'à agréger les dépenses au niveau de chaque GHM. Enfin, le coût d'un GHM est exprimé sur une base relative, en nombre de points ISA (Indice Synthétique d'Activité). Les établissements hospitaliers peuvent alors décrire la lourdeur économique et médicale des patients traités dans leur structure et évaluer leur niveau d'activité en valorisant leur case-mix par l'échelle de points ISA par GHM. » (Lenay, 2000).

**VALORISATION EN POINT ISA
DOCUMENT DE L'AGENCE TECHNIQUE DE L'INFORMATION HOSPITALIERE¹¹**

Chaque année dans MAHOS est intégrée une liste de lettres-clés permettant de valoriser en points ISA l'activité externe des établissements.

Cette activité est déclarée par les établissements par le biais de fichiers produits par le logiciel FICHSUP.

Depuis 2001, les établissements doivent fournir deux fichiers correspondants à cette activité externe :

- Un fichier concernant l'activité externe hors urgence
- Un fichier concernant l'activité des urgences.

Ces deux fichiers contiennent les mêmes informations et sont donc traités de la même manière pour aboutir à un nombre de points ISA. Le nombre de points correspondant à l'activité externe correspond à la somme des points ISA calculés sur ces deux fichiers. Avant 2001, un seul fichier était fourni.

Ces fichiers fournissent pour chaque lettre-clé la somme des coefficients facturés.

La valorisation en point ISA est donc obtenue en faisant la somme des produits pour chaque lettre clé du nombre de coefficients par le nombre points ISA correspondant à un coefficient lu dans la table de l'année voulue.

Le nombre de points ISA correspondant à un coefficient est obtenu, pour la totalité des lettres clés " Officielles " (publiée au Journal officiel), par la division du coût en centimes de la valeur officielle de lettre clé, par la valeur du point ISA en centimes dans l'Etude Nationale de Coût (coefficient de conversion).

Pour les activités de scanner et d'IRM la valorisation s'effectue de manière plus indirecte.

En effet, pour ces 2 types d'équipements, la valorisation ne peut pas s'effectuer directement à partir de la valeur d'une lettre clé, mais elle s'effectue à partir de la valeur d'un forfait qui dépend du type de machine ainsi que de son ancienneté ; ces valeurs de forfaits sont également publiées au JO. De plus, il existe pour chaque équipement une valeur de forfait normal et une valeur de forfait réduit, correspondant à la prise en compte de l'amortissement de l'équipement.

Pour effectuer cette valorisation nous avons créé des lettres clés particulières permettant de coder la valeur du forfait normal et réduit, ainsi que le nombre d'examen réalisés à la valeur de ces forfaits.

Ces lettres clés sont :

Pour le scanner :

Valeur du forfait normal : SCFNx

Valeur du forfait réduit : SCFRx

Nombre de scanner réalisés à forfait normal : SCNNx

Nombre de scanner réalisés à forfait réduit : SCNRx

Pour les IRM :

Valeur du forfait normal : RMFNx

Valeur du forfait réduit : RMFRx

Nombre d'IRM réalisées à forfait normal : RMNNx

Nombre d'IRM réalisées à forfait réduit : RMNRx

Le 'x' correspond à un identifiant permettant de saisir l'activité de plusieurs équipements de scanner ou d'IRM.

¹⁰ Depuis 1995, le ministère en charge de la Santé publie chaque année un BO présentant la méthodologie et les résultats de l'étude nationale des coûts par GHM. On trouve en particulier dans le fascicule tous les détails concernant la comptabilité analytique mise en œuvre, les unités d'œuvre permettant la ventilation des dépenses liées aux actes médico-techniques (Indices de coût relatif) et les méthodes de ventilation des coûts indirects. (note de l'auteur cité)

¹¹ Nous reproduisons ici à titre illustratif l'un des modes d'indexation relatif aux activités externes des hôpitaux, le MAHOS, mais pour une vision exhaustive de cette activité, il faudrait aussi reproduire d'autres modes de valorisation comme le MIGAC relatif à certaines activités internes.

Pour la valorisation en point ISA, le calcul est le suivant :

On multiplie la valeur du forfait normal de l'équipement n°1, par le nombre d'examen réalisés sur cet équipement à la valeur de ce forfait que l'on divise, comme pour les lettres clés standard, par la valeur du point ISA en centimes dans l'Etude Nationale de Coût (VCENC).

Par exemple pour un scanner portant le numéro 1 on obtiendrait pour ce qui est de son activité à forfait normal:

Valorisation en ISA de l'activité à forfait normal = $SCFN1 \times 100 \times SCNN1 / VCENC$
(à SCFN1 est en centimes)

Ce qui est fourni dans la table que vous pouvez consulter correspond au facteur $1/VCENC$, pour les lettres clés SCFNx, SCFRx, RMFNx, RMFRx.

Pour récupérer une table en format texte utiliser la commande d'enregistrement de votre navigateur pendant la visualisation de la table en format texte.

Dans une enquête relative à l'introduction de l'outil de gestion ISA dans le milieu hospitalier, Jean-Claude Moisdon (2000) a montré comment le point ISA devenait progressivement l'étalon de référence de la capacité de gestion des directeurs d'hôpitaux, au point qu'une donnée de base des échanges professionnels devenait la réponse à la question : « quelle est la valeur de ton point ISA ? » Passé en quelques années d'un simple instrument de connaissance des malades et de gestion des lits à un véritable instrument d'évaluation de la performance des hôpitaux, le PMSI a ainsi connu une série de développements plus ou moins inattendus impliquant des réorientations significatives de l'action gouvernementale. En effet, les agences régionales, sur la bases de modèles d'incitation, taxent aujourd'hui les établissements ayant un point ISA élevé, et restituent des moyens aux établissements sous-dotés. Bel exemple d'évolution des usages d'un instrument d'évaluation dans le sens d'un durcissement de la contrainte financière.

Les critiques adressées à cette démarche d'évaluation sont de deux ordres. Tout d'abord, l'établissement de valeurs de référence permettant l'indexation provoque une décontextualisation qui peut aboutir à des absurdités pratiques. Pour ce qui concerne l'évaluation de la performance des services ambulanciers en Grande-Bretagne, Chris Hood (2004) a bien montré les effets de standardisation des comportements provoqués par ce type d'indexation. Dans le cas qu'il étudiait, l'étalonnage sur une valeur moyenne de 8 minutes de trajet dans un périmètre donné avait eu pour conséquence que les ambulanciers, quelle que soit la distance du trajet, mettaient un point d'honneur à mettre 8 minutes, prenant leur temps pour les trajets courts, ou déployant un zèle énorme pour les trajets longs.

Dans un autre domaine d'activité, il peut ressortir que les critères d'évaluation de la qualité paysagère (variété des couleurs, lignes de fuites, etc.) s'appliquent très bien à une décharge de vieilles voitures (Charretton et Fischesser, 1995). On touche ici à une limite de la démarche instrumentale d'évaluation. Le pouvoir de conformation des activités sociales aux *optima* définis par l'instrument n'est jamais assuré, dans la mesure où les catégories prédéfinies subissent une série de traits qui peuvent être exemplifiés par une multitude d'objets ou de comportements.

Par ailleurs, il ne faut pas négliger, au delà du jeu des catégories, la capacité des individus à jouer avec les catégories, et à développer aussi ce que Michel de Certeau (1971) appelait des « pratiques rusées ». Pour ce qui concerne notre exemple de gestion des hôpitaux, le durcissement de l'indexation de l'allocation des ressources sur les indices affectés aux DRG aux Etats-Unis a eu des effets inattendus : les groupes professionnels visés ont configuré et diffusé des logiciels de gestion destinés à « optimiser le patient » au regard des règles de dotation : pour une série de symptômes donnés, le logiciel choisit la catégorie de pathologie la plus rentable.

Tous ces éléments permettent de comprendre les critiques les plus régulièrement adressées à cette démarche d'évaluation. Elle opère à l'évidence une réduction de la complexité des organisations ou des mondes sociaux. Le travail de sélection des informations pertinentes permet évidemment de

hiérarchiser les priorités d'action, mais il affecte durablement le sens de l'activité concernée. L'absurdité de certains résultats, comme la « très grande qualité paysagère » d'une décharge de voitures peut ainsi conduire à mettre en doute l'efficacité de cette démarche d'évaluation. Ces critiques, si elles sont valides, ne suffisent pourtant pas à disqualifier la démarche, sauf à s'en tenir à une posture anti-institutionnelle. On ne peut en effet occulter le fait que la critique en elle-même offre des alternatives et permet de reconfigurer les outils d'évaluation. De plus, l'intérêt de cette démarche, dans une perspective démocratique, tient dans le fait que les arbitrages politiques entre des voies de développement différentes sont nombreux. Hiérarchiser les mondes possibles et arbitrer les méthodes, même si cela impose une certaine technicité, ne relève pas forcément d'une logique de contrôle de type disciplinaire. La critique foucauldienne porte lorsqu'elle s'applique aux usages de cette démarche, mais elle s'applique difficilement aux opérations cognitives sur lesquelles repose la logique même de cette démarche. Ce qui pose problème en effet, c'est l'inclination inconsidérée au durcissement des systèmes d'indexation et le pilotage centralisé des outils de gestion.

3 LA DEMARCHE PRAXEOLOGIQUE

La notion de démarche praxéologique d'évaluation s'applique, dans l'acception que nous lui donnons, à tout travail de diagnostic reposant sur la compréhension des pratiques d'agents ou d'utilisateurs et/ou les interactions entre eux. Plusieurs méthodes peuvent être utilisées, de la statistique descriptive de type enquête sur les publics à l'observation participante (Weller, 2001 ; Dubois, 2001 ; Boussard, Mercier, Tripier, 2004). Quelle que soit la méthode employée, la démarche consiste d'abord à définir un ensemble de pratiques à analyser, à dresser ensuite un nombre plus ou moins important d'hypothèses, à recueillir des données sur cette base, puis à comparer des explications causales issues de l'enquête, et enfin à hiérarchiser ces explications sous la forme d'un cadrage théorique permettant de poser le diagnostic des causes de succès ou d'échec d'une politique. La prise en compte d'informations diverses sur les pratiques peut permettre de construire des « exemplaires »¹² qui révèlent des phénomènes que les hypothèses ne pouvaient prévoir et qu'il aurait été impossible d'identifier sans procéder au travail d'enquête (Passeron, 1998).

Un nombre considérable d'évaluations est effectué chaque année sur ces bases. Nous ne retiendrons qu'un exemple, qui nous paraît assez significatif de l'intérêt de cette démarche. Il concerne une enquête menée par Pierre Tripier et ses collaboratrices à propos de la mise en place d'une politique de tri sélectif dans une ville de la banlieue parisienne. Afin de comprendre les difficultés de mise en œuvre de cette politique, l'équipe de sociologues a suivi la tournée des rippers et enquêté sur le parcours des bacs de déchets. L'un des principaux résultats des observations et des entretiens réalisés avec les personnels a été le constat selon lequel les employés de l'entreprise de collecte à laquelle cette gestion avait été déléguée par la ville changeaient parfois leurs horaires de tournée sans prévenir les gardiens d'immeubles qui avaient pour mission d'acheminer les bacs. Les gardiens se trouvaient donc exclus du cercle de tri, et traînaient les pieds dans la mise en œuvre du dispositif. La principale préconisation du rapport remis à la ville commanditaire a été de réintégrer les gardiens dans ce cercle par des processus de motivation de type réunions, pots de fin d'année, etc.

Le plan du compte-rendu de cette enquête va nous permettre de progresser dans notre réflexion :

« Nous verrons dans le cas présenté ci-après comment les résultats décevants des opérations de tri sélectif dans les zones d'immeubles peuvent être analysés comme la rencontre, ratée, entre les

¹² Un exemplaire est « un cas probant qui procure un accès épistémique aux traits qu'il exemplifie. Un cas probant « révèle, dévoile, transmet des aspects de lui-même » (Elgin, 1992). Il peut ainsi revivifier l'évident (donné par le sens commun ou un corpus de textes), ou au contraire exemplifier des traits excessivement inapparents au regard des hypothèses d'arrière-plan. Sur l'usage de la notion d'exemplaire pour comprendre l'opération de construction typologique (Vion, 2004).

définitions des organisateurs du tri et celles des gardiens d'immeubles. Et nous verrons surtout que le malentendu se fait sur la place occupée par les gardiens, sur leur appartenance bafouée au "cercle de tri". Ce diagnostic du dysfonctionnement montre que les explications de premier niveau formulées par les acteurs de l'organisation ne sont pas justifiées : l'échec relatif du tri sélectif n'est pas dû, comme le laissent entendre certains, à une mauvaise communication des objectifs et des méthodes aux habitants. Un renforcement du dispositif d'information auprès de la population, pour renforcer les apprentissages et les changements des habitudes, ne serait pas la solution la plus pertinente. Nous verrons également que les diagnostics sociologiques classiques (analyse des relations de pouvoir, analyses de type culturel) ne permettent pas de remonter aux causes premières, causes que nous faisons apparaître à l'aide d'un autre modèle théorique." » (Boussard, Mercier, Tripier, 2004)

A la lecture de ce plan, nous voyons d'abord que la démarche praxéologique est envisagée comme un mode de réponse à des échecs difficilement compréhensibles, où la rationalité gestionnaire est mise en défaut. Le concept de « résultats décevants » a bien des affinités avec la démarche utilitariste, mais l'analyse des résultats en question n'est pas outillée. Ce n'est pas le problème. Un économiste comme Piketty aurait essayé de calculer les rendements par site pour identifier les points précis où la collecte a un rendement inférieur aux attentes (en fonction de variables démographiques et autres), puis aurait commandé un questionnaire aux habitants pour tester des variables afin d'optimiser le tri (surcroît d'information, politique d'incitation financière ou autres). Notre équipe part au contraire d'un constat du commanditaire pour définir son objet d'enquête. La seconde phrase permet de comprendre ce qui justifie ce type de démarche. Trivialement : les dysfonctionnements ne sont pas compréhensibles si l'on ne regarde pas de près ce qui se passe et qu'on n'écoute pas les gens. C'est donc l'enquête qui en révèle la nature. Il n'y a pas de schématisation pertinente a priori ! Puisque ce sont justement les catégories ou les pratiques sociales tacites qui posent problème, l'objet de l'enquête sera de les reformuler. Le social est affaire de prise en compte. Le terrain prime, il n'y a pas lieu de procéder par le haut et de hiérarchiser hors contexte. Lorsque Boussard, Tripier et Mercier disent que « un renforcement du dispositif d'information auprès de la population, pour renforcer les apprentissages et les changements des habitudes, ne serait pas la solution la plus pertinente », il est intéressant de noter qu'ils vont réfuter la démarche instrumentale par un argument instrumental : maintenant qu'on sait cela, on ne va quand même pas investir du temps et de l'argent dans l'instrumentation. On va plutôt traiter humainement les gardiens. En soi, la solution n'est pas moins rationnelle, dans la mesure où elle peut avoir un impact non négligeable sur le processus. Mais elle contraste avec les démarches précédentes dans la mesure où elle n'a pas pour objet de mettre en place les outils permettant de mieux mesurer les rendements ou de tester a priori l'efficacité de la variable « gardien d'immeubles ».

Cette démarche d'évaluation est donc, comme les précédentes, beaucoup plus un style de raisonnement qu'une simple méthode d'enquête. Il serait en effet réducteur de prétendre que la démarche praxéologique se caractérise par l'usage de méthodes qualitatives comme l'observation directe, les entretiens, *focus groups*, et autres. De nombreux usages de la statistique descriptive nous semblent aussi s'inscrire dans ce type de démarche. Lorsque Pierre Bourdieu et Alain Darbel travaillaient en relation avec le service d'études et de recherches du Ministère de la culture dans les années soixante, la préoccupation du ministre Malraux était de comprendre les causes de ce qui était alors perçu comme l'échec de la politique de démocratisation culturelle initiée par le ministère. L'enquête menée par Bourdieu et Darbel sur les musées européens avait pour ambition d'offrir des éléments d'explication. Même si l'on peut discuter de la nature véritablement évaluative de cette enquête, il est intéressant d'en retracer les opérations, qui sont courantes dans un certain nombre d'évaluations. La première étape est celle d'une enquête exploratoire basée sur des entretiens libres, destinée à définir les catégories pertinentes d'un questionnaire adressé à des usagers issus de divers milieux sociaux. La deuxième est celle d'un échantillonnage basé sur la fréquentation des musées étudiés. L'absence de statistiques mensuelles des visites dans certains musées et l'hétérogénéité de la compatibilité ont contraint à un travail de contrôle fastidieux : établissement de moyennes

journalières, saisonnières, annuelles, puis sélection de musées représentatifs des différents types de musées, puis constitution d'un panel représentatif d'usagers et choix aléatoire des personnes interrogées. Le texte du livre *L'amour de l'art*, dont est issu l'enquête, retrace l'ensemble des phases de traitement statistique des questionnaires : premier programme d'analyse factorielle sur 1000 questionnaires, enquêtes de vérification des dires des enquêtés sur leur temps de visite effective du musée, établissements de coefficients de correction sur cette base, puis enquête sur l'ensemble de l'échantillon, traitement par corrélations, etc. Ce qui est intéressant dans ce type d'usage de la statistique, c'est que la volonté de comprendre des pratiques suppose un travail de vérification par l'observation ou l'entretien. D'autre part, elle est mobilisée pour une description fine des pratiques, mais pas pour le travail de construction des mondes possibles à partir de cet état du monde présent. Une fois le constat posé d'une différence de pratiques selon les milieux sociaux, l'élaboration des possibles s'appuie sur une critique sociale de la monopolisation des instruments d'appropriation des biens culturels, pas sur la recherche de l'effet pur d'une variable. Comme l'a bien résumé Alain Desrosières (2003b) : « L'acteur principal de ce théâtre statistique reste le groupe social ou l'individu, ce qui distingue complètement ce théâtre de celui des méthodes économétriques ultérieures (régression logistique) qui envahiront ensuite la sociologie quantitative, et dans lesquelles les « protagonistes » (et sujets des verbes dans les phrases) sont des « variables », c'est à dire des entités abstraites construites sur le modèle ahistorique des sciences de la nature. C'est de ce point de vue que la sociologie empirique de Bourdieu diffère radicalement de celles d'autres courants pour lesquels la question centrale et quasiment unique est celle de « l'effet d'une variable sur une autre », question qu'il ne pose pas dans ces termes. (...) Ceci n'implique pas que Bourdieu s'interdise de parler en termes d'« effet d'une variable » (notamment la fameuse « CS du père »), mais il refuse de s'enfermer dans la recherche des « effets purs d'une variable », résultant de la distillation fractionnée opérée par les modèles économétriques de type LOGIT ».

Le résumé que propose Desrosières montre bien que les types de démarche ne se résument pas à un débat entre méthodes qualitatives et quantitatives mais bien à des styles de raisonnement qui peuvent cliver l'usage même d'un type de méthode.

4 LA DEMARCHE PROCEDURALE

Nous entendons par démarche procédurale toute démarche qui part du postulat de l'irréductibilité des valeurs pour définir des procédures de médiation ou de délibération susceptibles de faire émerger un accord entre parties intéressées à un enjeu commun. Comme nous nous centrons sur l'évaluation de l'action publique, nous excluons de notre réflexion toutes les questions désormais classiques relatives à la procéduralisation de la justice, pour nous centrer sur les procédures de débat public. Sous l'effet de la forte diffusion des travaux de Jürgen Habermas sur la démocratie délibérative, on ne compte plus les promoteurs ou les observateurs du tournant délibératif dans la conduite de l'action publique¹³. Ce qui nous intéresse plus particulièrement ici est l'usage proprement évaluatif qui peut être fait des procédures de débat comme les conférences de citoyens ou les procédures de débats publics les plus récemment organisés en France par la Commission Nationale du Débat Public (CNDP). Nous n'ignorons pas que les procédures de débat public sont bien plus que de simples exercices d'expertises¹⁴. Mais ce qui nous intéresse ici, c'est la tendance de plus en plus affirmée à l'organisation de grands débats régionaux ou nationaux destinés à évaluer *ex ante*, c'est-à-dire en amont de la décision, à un moment où l'opportunité des projets peut être discutée. Les récents débats organisés par la CNDP à propos de la ligne grande vitesse PACA ou de la gestion des déchets radioactifs nous semblent aller dans ce sens. Nous faisons donc l'hypothèse qu'émerge en France un type nouveau d'évaluation *ex ante*, portée par les acteurs eux-mêmes. Il

¹³ Pour faire le point sur ces auteurs et leurs débats, nous recommandons la lecture du numéro spécial « *Debate* » de la *Revue suisse de science politique*, paru en décembre 2004.

¹⁴ Pour une vue d'ensemble de la place et du rôle des débats dans le cours de la décision publique, on se reportera avec profit aux références suivantes : Lafaye, Thévenot, 1993 ; Pourchier, 1994 ; Charlier, 1999 ; Leborgne, 1999 ; Lolive, 1999 ; Blatrix, 2000 ; Cefai, Trom, 2001 ; Rui, 2001 ; Fourniau, 2001, 2004

s'agit alors de comprendre le cheminement par lequel est traité l'information dans ce type de démarche. Contrairement aux démarches précédentes, l'objet n'est pas seulement de sélectionner ou de recueillir des informations, mais de les mettre en commun et de faire émerger des arguments nouveaux de la délibération. Comme le résumait bien un président de commission de débat public : « ce que l'on cherche, ce n'est pas un sondage, ce sont des arguments pour et des arguments contre qui permettent de définir une solution qui convienne à tous »¹⁵.

La procédure de débat public institutionnalisé permet le libre accès de tous, chaque citoyen qui s'estime concerné peut participer au débat et contribuer à élargir l'éventail des mondes possibles (Fourniau, 2004). La nature des interventions au débat ne peut être anticipée et le caractère prévisible de certaines interventions ne doit pas occulter le fait que surgissent des arguments qui n'avaient pas été anticipés et qui sont pourtant techniquement recevables. Ce que soulignent les études sur la démocratie technique, c'est justement la capacité des profanes à formuler des avis scientifiquement recevables et pourtant non anticipés par les scientifiques (Callon, Barthe, Lascombes 2003). Dans notre vocabulaire, la prise en compte est ici maximale et toutes les garanties sont données pour qu'il en soit ainsi. L'opération de prise en compte est rendue transparente et, on serait tenté de dire, exemplaire : chaque intervention est transcrite et archivée. Les choses se compliquent singulièrement lorsqu'il s'agit d'effectuer une synthèse de ces contributions hétérogènes pour en dégager les faits marquants. Les opérations de hiérarchisation ne pourront être explicitées, quels que soient les efforts déployés en ce sens par la personne en charge de la rédaction finale.

La démarche est donc totalement opposée à la démarche utilitariste, où c'est l'exact opposé qui se produit : les informations prises en compte sont limitées et le mode de hiérarchisation est protocolarisé dans les moindres détails. Baliser le chemin qui mène des débats jusqu'à des recommandations pratiques est tout simplement impossible. Le récent débat sur l'école organisé par le gouvernement Raffarin en fournit un exemple frappant : même si un luxe de précaution est pris pour que la synthèse soit exemplaire (recours à des universitaires spécialisés dans le traitement informatique des données textuelles, présence d'un conseil scientifique prestigieux, mise en ligne d'un maximum d'information relative au traitement de l'information, etc), il reste délicat d'explicitier comment on passe de 15 000 débats et de centaines d'interventions écrites -et des mètres cubes de papier associés !- à un rapport de quelques centaines de pages. En d'autres termes, l'opération qui publicise cette construction relève d'un travail de synthèse dont il est difficile d'explicitier les règles propres, la manière dont la synthèse de ces débats a été effectuée par Claude Thélot ne peut être explicitée. Cette limite méthodologique justifie le plus souvent la prudence des conclusions des experts. Les mondes possibles sont moins strictement hiérarchisés que présentés sous forme d'options plus ou moins souhaitables au regard des opinions exprimées au cours des consultations.

Le rapport de M. Marzolf, président de la commission particulière du débat public sur la ligne à grande vitesse PACA, est tout à fait éclairant. Le début du rapport insiste sur la prise en compte, avec l'ensemble des moyens employés pour recueillir le maximum d'avis argumentés (voir encadré).

¹⁵ Philippe Marzolf, président de la commission particulière du débat public sur la ligne ferroviaire à grande vitesse PACA, interview TV à France 3 Provence, 9. 05. 2005.

MODALITÉS DE CONSULTATION DU DÉBAT PUBLIC SUR LA LGV PACA (SOURCE : RAPPORT MARZOLF)

- exposition itinérante sur le débat public :
 - 128 lieux d'exposition, 18 semaines d'exposition, 79 lieux de tractage
 - 39000 journaux et autant de synthèses du dossier du maître d'ouvrage tractés en marge de l'exposition (sur un total de 80000).
- un site Internet (189 connexions quotidiennes en moyenne, 35000 visites en 6 mois)
- une carte T pour s'abonner à l'actualité du débat : fichier de 9620 noms collectés. Une grande partie des 7500 dossiers du maître d'ouvrage diffusés l'ont été suite à une demande par carte T.
- un système de questions/réponses : 1200 questions posées (dont 277 avis)
- un journal du débat : 5 numéros, dont le premier, diffusé à 1 500 000 exemplaires, a présenté le débat, son calendrier, ses modalités. Il offrait la possibilité de s'abonner à l'actualité du débat et de recevoir la synthèse du maître d'ouvrage. Le journal n° 5 sera diffusé dans les premiers jours de septembre et portera à la connaissance du public la synthèse du compte rendu de la CPDP.
- 39 réunions publiques réparties comme suit :
 - 3 réunions de lancement, à Nice, Toulon et Marseille ;
 - 3 auditions publiques pour que les acteurs fassent connaître leur point de vue au début du débat ;
 - 8 réunions thématiques sur les grands thèmes que la commission avait identifiés au début du débat : Parc méditerranéen, les questions foncières et d'urbanisme, les enjeux d'environnement (2 réunions), le développement territorial, les transports régionaux (2 réunions) et le transport des marchandises ;
 - 5 ateliers pédagogiques : sur le financement, les solutions alternatives proposées par les acteurs et la présentation du dire d'expert commandé par la Commission nationale du débat public (3 réunions) ;
 - 17 réunions de proximité réparties sur le territoire concerné ;
 - 3 réunions de synthèse à Nice, Toulon et Marseille.

Près de 8000 personnes ont participé à ces réunions (dont près de 50% dans le Var). D'autre part, plus de 300 contributions, délibérations et motions ont été adressées à la CPDP (et mises en ligne), ainsi que 14 pétitions ayant recueilli au total plus de 26 000 signatures, présentant des avis diversifiés sur le projet. Certaines contributions ont fait l'objet de cahiers d'acteurs : en effet, 55 d'entre elles ont été sélectionnées pour leur pertinence et leur apport significatif au débat. Ces cahiers d'acteurs ont été édités en moyenne à 15 000 exemplaires chacun et c'est près de 750 000 exemplaires au total qui ont été distribués dans les réunions et adressés aux abonnés à l'actualité du débat. »

On voit bien que ce type de démarche peut permettre de faire émerger les formes d'un accord sur l'état du monde présent, et sur les dilemmes dont il est porteur. Mais pour étayer un travail de construction des mondes possibles plus ordonné, il faut recourir à des opérations cognitives susceptibles de définir les critères de fermeture des options : analyse des coûts, sondages d'opinion, etc. C'est bien ce qu'illustre bien la conclusion du rapport Marzolf sur le débat relatif à la LGV PACA : « A la question, posée au débat, de l'opportunité d'une ligne nouvelle à grande vitesse dans la région PACA, la réponse est complexe : une ligne nouvelle, oui mais pas nécessairement de bout en bout ni forcément à grande vitesse (bien que cette dernière soit ressentie comme une très forte nécessité dans les Alpes-Maritimes). Et, surtout, à condition que cette ligne nouvelle soit assortie d'un vrai programme de transports régionaux, qu'elle respecte l'environnement, que les gares soient au cœur des agglomérations, qu'elle s'intègre à un vrai plan d'aménagement du territoire, et qu'elle apporte un réel progrès au quotidien (report modal, gain en temps de parcours...) ! Par conséquent, avant de répondre à la question de l'opportunité d'une LGV, les participants au débat souhaitent que le maître d'ouvrage et les élus prennent en compte leurs priorités et répondent à leurs inquiétudes. »

CONCLUSION

Si l'on suit la réflexion de Goodman sur les manières de faire des mondes, on voit bien que ce qui est en jeu dans la discussion est bien la correction (*rightness*) des versions du monde proposées par l'évaluateur. Les tenants de la démarche utilitariste insistent sur la stabilité des catégories et les propriétés logiques de leur méthode de construction des mondes possibles. Les tenants des démarches praxéologique ou procédurale insistent au contraire sur le fait que la compréhension fine

des événements à partir d'une recherche approfondie d'indices et de cas exemplaires permet la reconception de l'état du monde présent. Ils jugent donc beaucoup trop sommaire la version qu'en proposent les économètres. Mais leur construction des mondes possibles ne tient pas sa force probante d'un raisonnement formellement logique. Elle la tient du fait que la version de l'état du monde présent trouve preneur au regard de l'expérience commune. Les critères de jugement de la correction des versions n'étant pas les mêmes, et n'intervenant pas sur les mêmes versions mobilisées dans le raisonnement évaluatif (monde présent, mondes possibles), la controverse entre les tenants de ces démarches est sans fin. Si nous admettons avec Jean-Claude Passeron qu'« un raisonnement sociologique ne peut être de part en part ni de bout en bout un raisonnement expérimental » (Passeron, 1991), ces controverses ont peu de chance de s'apaiser grâce à l'improbable apparition d'une méthode qui mette tout le monde d'accord.

Notre apport est de qualifier plus avant l'incompatibilité entre différents styles d'évaluation en pointant l'impossibilité de combiner deux opérations, la hiérarchisation et la prise en compte, de manière transparente. Il reste à préciser si cette impossibilité est une impossibilité pratique ou logique.

Si nous essayons de faire travailler le tableau proposé, nous comprenons que ces démarches ne sont pas totalement exclusives. Il est possible de parcourir le continuum séquentiellement : compréhension des pratiques, analyse des préférences, débat, instrumentation, et vice versa (avec itérations successives éventuellement). La tendance est aujourd'hui d'expérimenter des démarches mixtes.

- La combinaison des démarches de type utilitaire et instrumental nous semble caractériser l'ensemble des opérations regroupées aujourd'hui sous le vocable générique de « nouveau management public » (Pollitt & Bouckaert, 2000). Ces opérations ont des effets politiques et sociaux importants dans la mesure où les hiérarchies produites peuvent justifier des transferts d'allocation des ressources parfois assez radicaux¹⁶.
- A l'opposé, la combinaison d'enquêtes sur les pratiques et de débats publics provoque des changements beaucoup plus incrémentaux, et tend à reproduire ce que nous appellerions volontiers une « sphère de gestion concertée ».
- Une démarche de plus en plus fréquente tend à concilier les démarches situées au centre du tableau par le recours au benchmarking, qui consiste à produire des indicateurs et à mettre en œuvre des changements sur la base d'études de cas visant à l'identification des meilleures pratiques (*good practices*). L'Union Européenne fait un usage étendu de ce type de démarche.
- D'autres innovations plus radicales visent à concilier les démarches utilitaristes et procédurales. Le meilleur exemple est celui des sondages délibératifs mis en œuvre par James Fishkin et ses collaborateurs (Fishkin, Luskin, 1996 ; Fishkin, Luskin Jowell, 2003), qui consistent à alterner des phases de débats en conférences de citoyens et des sondages relativement élaborés. Cette démarche combinatoire permet en quelque sorte de tester les arguments décisifs et de réduire progressivement le champ des possibles.

Le choix d'une démarche est souvent suspecté de dissimuler des intérêts particuliers et d'inavouables idéologies. Les développements présentés ici n'épousent pas cette thèse souvent trop réductrice. Ils placent plus volontiers les antagonismes maintes fois constatés sur un plan logique ou pratique. Cependant, il est possible de montrer que privilégier un mode d'expertise revient à attribuer en creux une place spécifique au politique. On s'aperçoit que le rôle des autorités politiques dans la définition de la procédure d'évaluation est marginalisé aux deux pôles : celui où les mondes possibles sont fortement hiérarchisés selon un protocole scientifique, et celui où ces mondes possibles sont simplement élaborés par une opération de synthèse effectuée par une

¹⁶ L'exemple le plus paradigmatique étant sans doute le train des réformes thatchériennes, bien résumé par Patrick Le Galès (2004).

personnalité indépendante. On peut comprendre que les politiques qui souhaitent avoir prise sur le travail d'évaluation préfèrent les démarches proches des types instrumental ou praxéologique. A l'inverse, le pouvoir des évaluateurs, ou des savants pour reprendre l'appellation de Max Weber, est d'autant plus fort que l'on situe aux extrêmes du continuum. Le politique doit-il déléguer aux savants la charge de hiérarchiser les mondes possibles, quitte à fournir les critères de hiérarchisation ? Ou vaut-il mieux que l'expert se spécialise dans une fonction de prise en compte, ce qui le qualifie plutôt pour un rôle d'animateur du débat social ou de lanceur d'alerte ?

BIBLIOGRAPHIE

- Michel AMIOT, *Les sociologues contre l'Etat*, Paris, Ed. EHESS, 1986.
- Madeleine AKRICH, Michel CALLON, Bruno LATOUR, « A quoi tient le succès des innovations ? », *Annales des Mines*, 4 (29), 1988.
- Lusin BAGLA-GÖKALP, « Le chercheur et son instrument: changement des systèmes de mesure et des pratiques scientifiques en mécanique des fluides », *Revue Française de Sociologie*, XXXVII, 4, octobre-décembre 1996, pp. 537-566.
- Yannick BARTHE, Michel CALLON, Pierre LASCOUMES, *Agir dans un monde incertain. Essai sur la démocratie technique*, Paris, Seuil, Coll. La couleur des idées, 2001.
- Cécile BLATRIX, *La « démocratie participative », de Mai 1968 aux mobilisations anti-TGV. Processus de consolidation d'institutions sociales émergentes*, thèse de science politique, Université Paris I, janvier 2000.
- Pierre BOURDIEU et Alain DARBEL, *L'amour de l'art. Les musées et leur public*, Paris, Minuit, 1966.
- Valérie BOUSSARD, Delphine MERCIER, et Pierre TRIPIER, *L'Aveuglement organisationnel*, CNRS Editions, 2004
- Pierre CAHUC et André ZYLBERBERG, *Le chômage, fatalité ou nécessité ?* Paris, Flammarion, 2004
- Daniel CÉFAÏ et Danny TROM (dir.), *Les formes de l'action collective. Mobilisations dans des arènes publiques*, Editions de l'EHESS, Raisons pratiques, 2001
- Michel de CERTEAU, *L'invention du quotidien. 1. Arts de faire*, Folio essais, 1990.
- Philippe CHARRETTON et Bernard FISCHESSE (1995), "L'analyse paysagère dans la gestion des territoires." *Ingénieries - EAT* n°1: 31-40.
- Bruno CHARLIER, *La défense de l'environnement: entre espace et territoire. Géographie des conflits environnementaux déclenchés en France depuis 1974*, thèse de géographie et d'aménagement, Université de Pau et des Pays de l'Adour, 1999.
- Yves CROZET, « Calcul économique et démocratie : des certitudes technocratiques au tâtonnement politique », *Cahiers d'économie politique*. 2004. n°47. pp. 155-172.
- Alain DESROSIERES, *La politique des grands nombres. Histoire de la raison statistique*, Paris, La Découverte, 1993.
- Alain DESROSIERES, « Historiciser l'action publique : l'Etat, le marché et les statistiques », in LABORIER P. et TROM D., *Historicités de l'action publique*, Paris, PUF, 2003, pp. 207-221.
- Alain DESROSIERES, « Bourdieu et les statisticiens : une rencontre improbable et ses deux héritages », in ENCREVE B., LAGRAVE R.M., *Travailler avec Bourdieu*, Flammarion, Paris, 2003b, pp. 209-218.
- Patrice DURAN, Eric MONNIER, « Le développement de l'évaluation en France. Nécessités techniques et exigences politiques », *Revue française de science politique*, n°2, 1992, pp. 235-262.
- Patrice DURAN, « L'évaluation, au risque de son institutionnalisation », *Politiques et Management Public*, vol. 20, n° 4, décembre 2002, introduction (pp. I-V).
- James FISHKIN, Robert LUSKIN, « The deliberative poll. A reply to our critics », *The public perspective*, 1996, pp. 45-49.
- James FISHKIN, Robert LUSKIN, Roger JOWELL, "Considered opinions : deliberative polling in Britain", *British Journal of Political Science*, 32, 2003, pp. 455-487.
- Emmanuel FLACHAIRE et Guillaume HOLLARD, "Une approche comportementale de l'évaluation contingente". *Revue Economique*, à paraître, 2006.
- Joseph FONTAINE, « Evaluation et sciences sociales utiles. Raison des décideurs, offre d'expertise et usages sociaux dans quelques pratiques régionales », *Politix*, n° 36, pp. 51-71.
- Joseph FONTAINE, Philippe WARIN, « Retour d'évaluations : la régionalisation des politiques publiques entre affichage et incertitude », *Pôle sud*, n° 12, mai 2000, pp. 95-112.
- Jean-Michel FOURNIAU, « Mésentente et délibération dans les conflits d'aménagement : l'expérience du débat public institutionnalisé », in BLAIS, J-P., GILLIO, C. et ION J. (dir.), *Dynamiques associatives et cadre de vie*, Paris, Meltt-PUCA, 2001, pp. 261-301.
- François FOURQUET, *Les comptes de la puissance. Une histoire de la comptabilité nationale et du plan, encre, recherches*, 1980.

- Bastien FRANÇOIS et Érik NEVEU, *Espaces publics mosaïques. Acteurs, arènes et rhétoriques des débats publics contemporains*, Rennes, Presses de l'Université de Rennes, 1999.
- Ghislain GENIAUX, Guillaume HOLLARD ET Claude NAPOLEONE : « Peut-on quantifier le territoire ? » dans A. DONZEL (dir.), *Métropolisation, gouvernance et citoyenneté dans la région urbaine marseillaise*, MAISON-NEUVE & LAROSE, 2001
- Olivier GODARD (2004) "Autour des conflits à dimension environnementale - Evaluation économique et coordination dans un monde complexe", *Cahiers d'Economie Politique*, N° 47, pp. 127-153.
- Jürgen HABERMAS, *Droit et démocratie. Entre faits et normes*, trad. franç. R. Rochlitz et C. Bouchindhomme, Paris, Gallimard, NRF essais, 1997.
- Steve JACOB et Frédéric VARONE, *Evaluer l'action publique. Etat des lieux et perspectives en Belgique*, Academia Press, 2003
- Steve JACOB et Frédéric VARONE, « Cheminement institutionnel de l'évaluation des politiques publiques en France, en Suisse et aux Pays-Bas (1970-2003) », *Politiques et Management Public*, 22 (2), pp. 135-152.
- Steve JACOB et Frédéric VARONE, « Institutionnalisation de l'évaluation et nouvelle gestion publique : un état des lieux comparatif », *Revue Internationale de Politique Comparée*, 11 (2), pp. 271-292.
- Marie-Christine KESSLER, Pierre LASCOUMES, Michel SETBON, Jean-Claude THOENIG, (sous la direction de), *Evaluation des politiques publiques*, Paris, L'Harmattan, 1998
- Claudette LAFAYE et Laurent THEVENOT, « Une justification écologique ? Conflits dans l'aménagement de la nature », *Revue française de sociologie*, vol. 34, n° 4, Paris, 1993, pp. 495-524.
- Pierre LASCOUMES, « Pratiques et modèles de l'évaluation », in KESSLER M.-C., LASCOUMES P., SETBON M., THOENIG J.-C., *Evaluation des politiques publiques*, L'Harmattan, Coll. Logiques politiques, 1998.
- Pierre LASCOUMES, Michel SETBON, *L'évaluation pluraliste des politiques publiques : enjeux, pratiques, produits*, GAPP-CNRS-Commissariat Général au Plan, 1995.
- Pierre LASCOUMES, Patrick LE GALES, *Gouverner par les instruments*, Presses de Sciences PO, Coll. Gouvernances, 2004.
- Bruno LATOUR, *Politiques de la nature. Comment faire entrer la science en démocratie*, Paris, La Découverte, Armillaire, 1999.
- Mathieu LEBORGNE, *De quelques caractères associatifs. La mobilisation associative entre l'élu local et le citoyen lors du débat public relatif au projet de ligne THT entre Boute (Var) et Carros (Alpes-Maritimes)*, Ministère de l'Environnement et de l'Aménagement du territoire et SHADYC (EHESS), Marseille, janvier 1999.
- Olivier LENAY, « De la concrétisation d'un outil de gestion : le rôle du PMSI dans la régulation du système hospitalier », *Politiques et Management Public*, vol 18, n° 3, 2000.
- Jacques LOLIVE, *Les contestations du TGV Méditerranée*, L'Harmattan, 1999.
- Yves MENY, Jean-Claude THOENIG, *Politiques publiques*, Paris, PUF, 1989.
- Jean-Claude MOISDON, « Quelle est la valeur de ton point ISA ? Nouveaux outils de gestion et de régulation dans le système hospitalier français », *Sociologie du travail*, 42, 1, 2000, pp. 31-49.
- Jean-Pierre NICOLAS, Fabien DUPREZ, Sandrine DURAND, Fabrice POISSON, Pierre-Louis AUBERT, Mireille CHIRON, Yves CROZET, Jacques LAMBERT, *10 ans de monétarisation des effets locaux de la pollution atmosphérique*, Synthèse Inrets n°46, 2002
- Frédérique PALLEZ, « De la mesure dans un service public régalién. Peut-on et faut-il quantifier la charge de travail des magistrats ? », *Politiques et Management Public*, vol. 18, n°4, décembre 2000, pp. 91-118.
- Jean-Claude PASSERON, *Le raisonnement sociologique*, Paris, Nathan, 1991.
- Jean-Claude PASSERON, « Statique et dynamique. Caractériser, expliquer, comprendre le changement », *Revue européenne des sciences sociales*, Tome XXXVI, n° 110, 1998, pp. 199-263.
- J.-C. PASSERON, Louis-André GERARD-VARET, (dir.) *Le modèle et l'enquête. Les usages du principe de rationalité en sciences sociales*, EHESS, 1995.
- Thomas PIKETTY, « L'impact de la taille des classes et de la ségrégation sociale sur la réussite scolaire dans les écoles françaises : une estimation à partir du panel primaire 1997 », 2004, 71 p.
- Christopher POLLITT et Geert BOUCKAERT, *Public management reform. A comparative analysis*, Oxford University Press, 2000.

Code de champ modifié

- Éric POURCHIER, « De la CARDE à la FARE-Sud : le mouvement de contestation du TGV Méditerranée », in Alain HAYOT (dir.), *Ville et citoyenneté*, Villes et territoires méditerranéens, Montpellier, 1994.
- Sandrine RUI, *La démocratie en débat. Les citoyens face à l'action publique*, Paris, Armand Colin, 2004.
- Sandrine RUI, *L'expérience démocratique. Observation et analyse de l'implication des citoyens dans les procédures de concertation relatives à la mise en œuvre de projets d'infrastructure et d'aménagement*, thèse de sociologie, Université de Bordeaux II, janvier 2001.
- Vincent SPENLEHAUER, *L'évaluation de politiques, usages sociaux. Trois études de cas d'évaluation*, L'Harmattan, 1995.
- Vincent SPENLEHAUER, « Quel développement, quelle professionnalisation de l'évaluation ? », *Pouvoirs locaux*, n° 38, 1998.
- Vincent SPENLEHAUER, « Intelligence gouvernementale et sciences sociales », *Politix*, n° 48, décembre 1999, pp. 95-128.
- Vincent SPENLEHAUER, « L'évaluation au service des conseils régionaux », *Sociologie du travail*, n° 42, 2000, pp. 245-263.
- Jean-Claude THOENIG, « L'évaluation en actes. Leçons et perspectives », *Politiques et Management Public*, vol 20, n°4, 2002, pp. 33-50.
- Bruno VAN DER LINDEN, « L'effet des formations professionnelles de chômeurs : de l'impact sur les individus à l'impact macroéconomique », *Bulletin de l'IREES*, Bruxelles, n° 227, 2001 (<http://www.ires.ucl.ac.be>)
- Antoine VION, « Construire une typologie de *policy takers* : une opération logique et heuristique », *Revue Internationale de Politique Comparée*, 4^e trimestre, décembre 2004.