

## Congrès AFSP 2009

### ST 29 Pour une analyse des politiques publiques de l'éducation

Session n°1, « Sciences politiques et sciences d'État : l'ouverture problématique du secteur de l'éducation »

Xavier Pons (Largotec, Paris 12, [ponsx@wanadoo.fr](mailto:ponsx@wanadoo.fr))

#### **Morphologies et influences des sciences d'État en éducation. L'exemple de l'évaluation de l'action éducative en France**

À chaque rentrée scolaire, nous assistons en France à une inflation de discours sur le prétendu déclin de l'école française, inflation dont les éditeurs ont visiblement compris les bénéfices potentiels. En analysant le contenu d'une cinquantaine d'« ouvrages de controverse » recensés à cette rubrique dans la base « OPALE » de la Bibliothèque nationale de France en 1995, Vogler (1996) montre que la publication de ces derniers est globalement croissante depuis le début des années 1980 (avec un pic en 1984) et qu'ils partagent des caractéristiques communes : un ton virulent, le profil de leur auteur (professeurs, journalistes et essayistes, philosophes et pédagogues) et un discours catastrophiste sur l'état de l'école en France. Dans son analyse lexicale de onze ouvrages à forte diffusion publiés entre 1984 et 1997, Paveau (1999) va plus loin et parle d'une véritable *doxa* pour qualifier ce phénomène, d'un discours collectif reposant sur un nombre limité de métaphores et de mythes : la maladie du grand corps qu'est l'école, sa décadence, son âge d'or « des préaux et blouses grises » (p. 22) ou encore le complot des pédagogues contre les enfants du peuple.

Serait-il donc impossible en France d'avoir un débat rationnel sur le fonctionnement du système éducatif ? Qui produit les connaissances politiquement légitimes sur le fonctionnement, réel ou souhaitable, du système éducatif français et sur les réformes à entreprendre ? Pour répondre à ces questions, cette communication analyse l'évolution d'un domaine d'action publique en particulier, qui s'est grandement développé à partir du milieu des années 1970 : l'évaluation de l'action éducative<sup>1</sup>. Si l'évaluation, par ailleurs inégalement institutionnalisée en France selon les époques et les territoires (Dutercq, 2000, Pons, 2006, 2008), n'est bien évidemment pas le seul espace de production de connaissances sur le système éducatif, elle constitue un terrain d'investigation privilégié pour plusieurs raisons. Dès ses prémices, l'évaluation des politiques publiques a été présentée par ses défenseurs comme un outil d'objectivation scientifique des pratiques de gouvernement et de leurs effets (Nioche, Poinard,

---

1 Par « action éducative », nous désignons sous une forme condensée l'action publique en matière d'éducation. Il ne s'agit donc pas de l'acte éducatif et/ou pédagogique accompli dans la classe mais bien de la conception et la mise en œuvre de politiques publiques (quand elles existent), de mesures, de dispositifs ou de programmes d'action en tout genre. L'action éducative ne se limite pas aux réformes et aménagements gestionnaires du « système éducatif » (considéré ici comme la désignation particulière et historiquement situé de l'objet de la politique éducative), même si l'évaluation du système éducatif est la forme historique principale prise par l'évaluation de l'action éducative ces dernières années (Pons, 2008). Par convention, nous emploierons par la suite uniquement le terme d'évaluation, sauf dans les cas où celle-ci ne désignerait pas l'évaluation de l'action éducative.

1984, Monnier, 1987, Perret, 2003). Les chercheurs en science politique, en général sceptiques quant aux effets de l'évaluation sur la décision politique et la recombinaison des actions entreprises (CURAPP, 1993, Kessler, Lascoumes, Setbon, Thoenig, 1998), reconnaissent pour la majorité d'entre-eux que l'évaluation, tant dans sa dimension gestionnaire que dans une perspective démocratique (Duran, Monnier, 1992), contribue à rationaliser les actions publiques mises en œuvre et à les légitimer au regard de différentes valeurs (pertinence, efficacité, efficience etc.). Dans le secteur de l'éducation, l'apport décisif du développement de l'évaluation au sein du ministère en matière de production de données et de connaissances utiles à la description et la compréhension du fonctionnement du système éducatif est souligné aussi bien par les acteurs de ce système (Pair, 2001, DEP, 2004, Forestier, Thélot, Emin, 2007) que par certaines recherches, notamment en matière de production statistique (Pons, 2009).

La thèse que nous souhaitons mettre à l'épreuve est la suivante : dans le domaine de l'évaluation, nous assistons à la reconversion de différentes sciences d'État, sciences qui constituent la principale source de connaissance politique disponible actuellement pour penser le fonctionnement global du système éducatif. Par sciences d'État, nous désignons un ensemble de savoirs spécialisés utiles aux décideurs, à mi-chemin entre des pratiques administratives et des activités académiques, et visant à mettre en évidence des régularités dans les comportements des acteurs du système éducatif afin de mieux les gouverner (Ihl, Kaluszynski, Pollet 2003). Institutionnalisés de façon concomitante à la construction des États modernes, ces savoirs envisagent la société du point de vue de l'État, d'« en haut », et non en fonction des systèmes sociaux particuliers sur lesquels il entend agir. Ils conçoivent la société comme une série de questions ou problèmes à résoudre (la question sociale, la question urbaine, la question européenne etc.) et sont naturellement tournés vers l'organisation interne de l'appareil étatique, conçu lui-même comme le principal instrument de l'action publique (Leca, 1993).

Pour cela, nous procédons en trois temps. Sur la base de matériaux accumulés lors d'une recherche européenne sur le rôle des connaissances dans la régulation des systèmes éducatifs (Pons, van Zanten 2008)<sup>2</sup>, nous proposons tout d'abord une cartographie des producteurs de connaissance dans le domaine de l'évaluation au niveau national. Celle-ci met en évidence l'importance prise dans l'expertise publique en matière d'éducation par deux acteurs étatiques en particulier (inspecteurs généraux et statisticiens du ministère), dont les productions de connaissance reprennent explicitement les catégories d'analyse propres au fonctionnement interne de l'appareil étatique. Puis, à partir des données recueillies dans le cadre d'une thèse de doctorat (Pons, 2008), nous analysons les méthodes d'évaluation utilisées par ces deux acteurs. La recherche révèle que celles-ci ne changent pas de façon significative par rapport à ce que peuvent produire ces acteurs, soit dans d'autres cadres, soit à d'autres époques. Il existe des routines cognitives propres à chaque acteur. Nous en déduisons l'existence de deux sciences d'État en particulier : *l'art de l'extrapolation empirique* des inspecteurs généraux et la *science du constat chiffré* des statisticiens de la DEPP. Enfin, nous analysons l'influence exercée par ces dernières sur la politique d'évaluation du ministère de l'Éducation nationale depuis le début des années 1970 et, notamment, le statut qu'elles accordent aux travaux effectués par d'autres acteurs, en particulier ceux des chercheurs en éducation.

### ***Une cartographie des producteurs de connaissance dans le domaine de l'évaluation***

Cette cartographie (*mapping*) a été réalisée en 2007 par plusieurs chercheurs<sup>3</sup> sous la coordination d'Agnès van Zanten et moi-même. Elle constituait le premier axe d'une recherche financée par la Commission européenne et coordonnée par l'Université Catholique de Louvain -la recherche *Know&Pol*- visant à analyser le rôle de la connaissance dans la régulation des systèmes de santé et d'éducation<sup>4</sup>.

---

2 Il s'agit plus précisément d'une centaine d'entretiens auprès des évaluateurs et des destinataires de l'évaluation, de la constitution d'une statistique descriptive fondée sur une lecture systématique des publications, de l'administration d'un questionnaire et l'exploitation de biographies (pour la sociologie des membres de ces groupes) et enfin de la consultation d'une abondante littérature interne à ces organismes et de quelques archives (personnelles, officielles ou orales).

3 Choukri Ben Ayed, Hélène Buisson-Fenet, Roger-François Gauthier, Martine Kherroubi, Nathalie Mons, Elena Roussier-Fusco et Eric Verdier.

4 Une description précise du projet et l'ensemble des publications intermédiaires produites à ce jour sont disponibles en ligne,

L'objectif était tout d'abord de renseigner, grâce à des recherches documentaires, des questionnaires et des entretiens semi-directifs, une grille d'indicateurs quantitatifs et qualitatifs élaborée par une équipe de chercheurs internationale interne au consortium *Know&Pol* et destinée à permettre une première comparaison entre les pays participants au projet (France, Belgique, Écosse, Hongrie, Portugal et Roumanie). Il s'agissait ensuite de proposer, sur la base de cette grille, une première analyse synthétique de la production et de la circulation de connaissances dans le domaine étudié. La grille d'indicateurs invitait à analyser aussi bien certaines caractéristiques sociales ou morphologiques des acteurs concernés (*social mapping*) que les propriétés cognitives de leurs travaux (*cognitive mapping*). Concernant le rapport de synthèse, les coordonnateurs ont proposé une structure d'ensemble invitant, après avoir donné des éléments de contexte historiques et institutionnels, à distinguer les effets des types de structure, des types d'acteurs et des types de connaissances produites sur la circulation des connaissances elles-mêmes. Le tableau 1 mentionne, en anglais, les indicateurs proposés et leur description générale.

**Tableau 1. Indicators grid of Orientation 1 of the *KnowandPol* research**

<i>Indicator</i>	<i>General description</i>
<b>1) Formal policy making bodies that produce and use knowledge</b>	Listed in terms of permanent roles or office-holders
<b>2) Formal qualifications of office holders</b>	Detailed discipline and level of qualification
<b>3) Professional experience/careers of the decision makers</b>	Type and length of career, elected function (Yes or No), sector of activity, seniority, location of activity (National and/or Regional)
<b>4) Volume of knowledge production by body</b>	Total production, analysed by category: activity reports and/or participation to research/study (academic publication: Yes or No)
<b>5) Organisational tasks/responsibilities</b>	Definitions of tasks: leader, follower, advisory role, expert... Whether the body/policymaker obtains the knowledge from another body or whether it produces it itself
<b>6) Categories of knowledge production by use of knowledge</b>	For example: dissemination of knowledge; synthesis of knowledge; using knowledge to make a case/argument; seeking knowledge. Does the body/policymaker obtain/commission the knowledge construction/production from another source?
<b>7) Cognitive categories used to describe the sector/system</b>	Picture of sector, priorities, issues and problems (e.g. efficiency, economy, skills, reducing budgets)
<b>8) Budget dedicated to research</b>	Proportion to the total budget
<b>9) Research priorities</b>	Areas selected for funding. How are those priorities determined (Scientific committee, administrative body, etc.) – any political priorities?
<b>10) Who is funded</b>	Organisations selected for funding: public/private? Research/consultancy? National/local/international? Is the selection process public and transparent?
<b>11) Formal Relationships between the body and others</b>	Does the body have specific relationships with others in terms of knowledge? Are there specific instances or places for these relationships (e.g. joint membership of committees)?
<b>12) Use of international knowledge based categories</b>	Use of some indicators, documents, categories, discourses, words from international bodies
<b>13) Knowledge Producers</b>	Number, diploma, profession, scientific discipline
<b>14) Typology of the knowledge producers according to their relations with the policy makers</b>	By externality, association, clients, integrated, partners...

Source : Pons, van Zanten (2008, p. 72).

L'équipe de recherche s'est focalisé sur les producteurs de connaissance intervenant au niveau national dans la production de connaissances en matière d'évaluation et d'expertise du fonctionnement du système éducatif. Certains ont une mission officielle d'évaluation et sont facilement repérables : il s'agit des

inspections générales et de la DEPP<sup>5</sup>. Pour les autres acteurs sans mandat explicite, que nous appellerons les évaluateurs latents, trois critères ont présidé à leur identification et sélection : l'emploi du terme « évaluation » pour décrire les tâches de l'organisme, que ce soit dans des publications ou au cours des entretiens avec certains de ses membres, l'existence de liens formels ou cognitifs entre évaluateurs officiels et évaluateurs latents, et la formulation, dans les travaux des acteurs concernés, de jugements évaluatifs<sup>6</sup> sur le fonctionnement du système éducatif.

Cette cartographie met en évidence le poids des acteurs étatiques dans le domaine de l'évaluation, par ailleurs les seuls à avoir un mandat officiel d'évaluation. Aux côtés des syndicats d'enseignants et de chefs d'établissement, on dénombre peu d'instituts de recherche dont la production ne se réduit pas à des initiatives individuelles et qui produisent des connaissances sur le système éducatif de manière continue et cumulative. Trois se distinguent en particulier (l'Iredu<sup>7</sup>, l'INRP<sup>8</sup> et le Céreq<sup>9</sup>). Par ailleurs, on compte plusieurs acteurs absents, dont la production est soit nulle, soit peu significative, comme les cabinets de consultant privés (en tout cas au niveau national), l'UNESCO, l'Union Européenne, le Conseil d'analyse stratégique ou les associations de parents d'élèves. En revanche, les acteurs placés directement sous la tutelle du ministre (inspections générales, DEPP, HCEE<sup>10</sup> mais aussi INRP et Céreq) ou directement liés à l'État et à ses élites (Cour des comptes, inspection générale des Finances) sont bien plus représentés.

Il en va de même en termes de volume de publications et d'effectifs mobilisables. Depuis 1989, les inspections générales, qui comptent selon les années entre 240 et 350 membres, publient (soit dans leur rapport annuel synthétique, soit dans les rapports thématiques en ligne) entre 16 et 49 rapports par an (à l'exception de l'année 1998, marquée par le ministère de Claude Allègre). D'après la base documentaire en ligne de la DEPP (base Acadoc), cette direction d'administration centrale, composée d'entre 211 et 253 agents selon les années depuis 1992, a publié entre 1992 et 2007 entre 10 et 25 études statistiques à dimension évaluative. Aucun autre acteur ne produit autant d'évaluations sur le système éducatif français, sous quelque forme que ce soit. Même la trentaine de membres de l'Iredu, dont « l'évaluation du fonctionnement des institutions éducatives » constitue l'un des axes de recherche à part entière à partir de 1990, n'atteint pas un tel ordre de grandeur<sup>11</sup>.

La lecture des enquêtes et rapports publics, nécessaire pour le volet plus cognitif de cette cartographie (voir notamment les indicateurs 7 et 12 mentionnés dans le tableau ci-dessus), révèle quant à elle que ces deux acteurs dominants usent principalement de catégories descriptives renvoyant à l'architecture institutionnelle et au fonctionnement interne du système éducatif (catégories de personnels, ordres d'enseignements, sigles en tout genre etc.). Sur quelles méthodes d'analyse ces productions de connaissance reposent-elles ?

### ***Méthodes et routines cognitives<sup>12</sup> des évaluateurs officiels***

Concernant les inspecteurs généraux, l'analyse de l'ensemble des rapports publics met en évidence trois faits principaux. Premièrement, pratiquement tous les rapports sont fondés sur l'usage de trois techniques d'enquête en particulier : l'amas et l'analyse de données secondaires, l'observation sur le

---

5 Direction de l'évaluation, de la prospective et de la performance du ministère de l'Éducation nationale.

6 Par ce terme, nous désignons de façon générique tout jugement de valeur, tout propos normatif sur le fonctionnement souhaitable du système éducatif ou toute activité de formulation de propositions ou de recommandations.

7 Institut de recherche en éducation.

8 Institut national de la recherche pédagogique.

9 Centre d'études et de recherche sur les qualifications.

10 Haut conseil de l'évaluation de l'école, en activité de 2000 à 2005.

11 Le rapport scientifique de l'Iredu de l'année 1989-1990 recense certes près de 23 travaux d'évaluation, mais ceux-ci concernent tous les ordres d'enseignement (y compris l'enseignement supérieur) et différents pays (notamment les pays en voie de développement).

12 Par routines cognitives, nous désignons le ou les principes à l'origine de l'ensemble des choix méthodologiques effectués par un groupe professionnel, de façon plus ou moins consciente mais le plus souvent sur le mode de l'habitude, au moment où il est supposé produire le savoir spécialisé qu'on attend de lui.

terrain, le plus souvent en accord avec les autorités académiques, et des entretiens sur place. En général descriptifs et littéraires, les rapports d'inspection générale consistent ainsi généralement en une description des faits observés, assortie de recommandations, le passage des observations aux recommandations étant le plus souvent discrétionnaire.

Deuxièmement, ce savoir-faire est rationalisé par l'usage de multiples techniques d'enquêtes qui visent à améliorer la scientificité, ou en tout cas la portée générale, des conclusions : échantillonnage préalable aux visites, techniques permettant de chiffrer les phénomènes observés (exploitation, voire production de statistiques, administration de questionnaires, dénombrements localisés) et, de façon plus ponctuelle, techniques permettant de classer les phénomènes observés (construction de typologies empiriques, premières comparaisons).

Troisièmement, ce mouvement de rationalisation n'empêche pas la permanence de plusieurs biais dans l'analyse, biais qui sont à la fois connus et ignorés. Par exemple, la pertinence de l'objet d'évaluation est rarement interrogée, la possible dépendance à l'égard des sources secondaires connue, l'usage des guides méthodologiques limité. De même, les échantillonnages sont rarement justifiés et principalement pensés comme devant permettre une observation de diversité, ce qui pose deux problèmes principaux. D'une part, les inspecteurs ont tendance à accorder un privilège aux unités dans lesquelles les dernières actions sont effectivement mises en œuvre. D'autre part, la couverture territoriale permise par ce type d'échantillonnage est très inégale, certaines académies étant systématiquement plus visitées que d'autres (Pons, 2008).

Pour les inspecteurs généraux, le passage à l'évaluation ne se traduit donc pas par un changement radical de leurs méthodes, mais par la reconversion d'un ancien savoir-faire dont ils ont eu historiquement le monopole (l'inspection générale dans les établissements). Le principe à l'origine de leurs choix méthodologiques reste le même, il s'agit avant tout de vérifier sur le terrain.

En ce qui concerne les statisticiens du ministère, l'analyse de l'ensemble des enquêtes publiques disponibles en ligne sur la base documentaire de la DEPP (base Acadoc) à la rubrique « évaluation » met elle-aussi en évidence quatre faits principaux. Tout d'abord, l'on observe un déséquilibre important entre l'usage régulier et conséquent de n'importe quel type d'indicateur chiffré à vocation descriptive (nombres totaux, écarts-types, taux etc.) et un usage irrégulier et beaucoup moins répandu de techniques statistiques utilisées dans une perspective explicative (modélisation, analyses de régression, construction d'indicateurs *ad hoc* etc.). L'on observe également une volonté croissante de classer les phénomènes observés qui se manifestent par exemple par la construction presque systématique de quantiles, l'administration régulière de questionnaires, ou la construction de typologies.

Par ailleurs, les travaux de la DEPP donnent très rarement lieu à la production de jugements évaluatifs (20 cas répertoriés sur 429 travaux évaluatifs recensés par la base entre 1968 et 2007). Enfin, la production de travaux statistiques à dimension évaluative s'accompagnent parfois de véritables paris sur la méthode à privilégier. C'est le cas par exemple quand les statisticiens préfèrent une exploitation artisanale des réponses aux questions ouvertes d'un questionnaire à l'analyse textuelle ou lexicométrique par un logiciel. C'est le cas également quand ils acceptent de dépendre de plus en plus des fichiers de gestion pour construire des bases de données ou récolter les informations initiales pour un nouveau panel, ou quand ils décident de passer d'interrogations en face-à-face à des questionnaires téléphoniques. Ces choix méthodologiques résultent d'une appréciation qui ne repose pas seulement sur des critères scientifiques, mais aussi sur la plus-value escomptée de l'organisation d'une évaluation au regard d'une expérience professionnelle et des ressources disponibles à un moment donné (à la fois financières, techniques et organisationnelles).

Là encore, les méthodes utilisées par les statisticiens dans leurs travaux d'évaluation ne diffèrent pas de manière significative de celles qui peuvent (ou ont pu) être utilisées dans le cadre d'enquêtes statistiques générales. Le principe à l'origine de leurs choix méthodologiques reste le même, il s'agit avant

tout de décrire et classer les phénomènes observés.

La permanence de méthodes qui évoluent peu dans le temps, la persistance de biais méthodologiques à la fois connus et ignorés, l'orientation de ces méthodes vers l'action et la prise de décision tout comme l'usage de catégories d'analyse principalement descriptives et concentrées sur l'organisation interne de l'appareil étatique nous amènent à conclure à l'existence de sciences d'État en éducation, sciences qui perdurent malgré les injonctions au changement véhiculées par un instrument d'action publique *a priori* typique de la nouvelle gestion publique. Il s'agit ici de *l'art de l'extrapolation empirique* pour les inspecteurs généraux et de la *science du constat chiffré* pour les statisticiens de la DEPP. Quelle est l'influence de ces sciences d'État sur la politique d'évaluation de l'action éducative en France ?

### ***Sciences d'État et politique d'évaluation***

Du fait des incertitudes entourant la demande politique d'évaluation par les décideurs et les acteurs du système éducatif et d'une institutionnalisation inachevée de cette dernière en France, l'évolution de la politique d'évaluation de l'action éducative s'explique principalement par la dynamique de la concurrence que se livrent les évaluateurs entre eux (Pons, 2008). Cette concurrence est une concurrence cognitive, entre différentes sciences de gouvernement (Ihl, Kaluszynski, Pollet, 2003). Elle a une dynamique propre : il s'agit en clair d'une course à la généralité scientifique utile aux décideurs. Les acteurs recherchent d'abord des complémentarités avec les approches de leurs concurrents. Ils approfondissent ainsi leur connaissance de ces derniers (Simmel, 1992) et tentent ensuite de proposer une science de gouvernement de portée plus générale, intégrant toutes les autres comme autant d'éléments particuliers, partiels ou inachevés (processus de subsumption envisagé par Abbott, 1988). Cette concurrence se fonde enfin sur trois clivages principaux permettant de distinguer les productions de connaissance (rapports, enquêtes, conférences etc.) entre elles : l'opposition entre approches quantitatives et qualitatives, le conflit de légitimité entre observation interne et observation externe, et la différence entre analyse localisée et regard synoptique.

Jusqu'en 1992, *l'art de l'extrapolation empirique* constitue la science de gouvernement dont la portée est la plus générale et ce, pour plusieurs raisons. Jusqu'en 1984, la politique d'évaluation consiste avant tout à évaluer les élèves, soit de façon individualisée (en classe, par le biais des enseignants dont on contrôle l'exactitude de l'enseignement et des méthodes pédagogiques d'évaluation par le biais de corps d'inspection) soit de façon standardisée avec les évaluations internationales de l'IEA (fondée en 1958) qui se développent à partir des années 1960 ou les évaluations sur échantillon proposées par le service statistique du ministère. L'enjeu politique majeur pour les professionnels de l'évaluation est alors d'être capable, sur la base de savoir et savoir-faire différents, de délivrer un constat de portée générale sur l'état des connaissances et des résultats des élèves, dans un système scolaire français encore largement construit autour de filières pédagogiques verticales, par disciplines, types d'établissement ou ordres d'enseignement. Deux visions s'opposent clairement pendant cette période, qui renvoient à deux (puis trois) groupes professionnels différents. D'un côté l'évaluation est comprise comme une pratique traditionnelle d'enseignement, donc comme un art ancestral que seuls les praticiens de la pédagogie que sont les enseignants et leur hiérarchie (corps d'inspection) maîtrisent et peuvent transmettre. De l'autre, l'évaluation renvoie à un courant de recherche en expansion, plutôt mal perçu en France à cette époque, fondé sur des techniques psychométriques et développé par des experts étrangers et quelques chercheurs français isolés.

Dans le contexte scientifique de l'après 1968 (faible développement de la psychométrie, institutionnalisation récente et brutale des sciences de l'éducation, hésitations du service statistique ministériel, résistances parfois idéologiques du milieu des enseignants chercheurs au développement de tests standardisés), la concurrence cognitive est alors clairement à l'avantage des inspecteurs généraux. Ceux-ci sont en effet dans une position institutionnelle (sommet de la hiérarchie pédagogique du ministère) qui leur permet de transformer des expériences personnelles en discours pédagogiques de

portée générale, tout en critiquant le « tout quantitatif » de certains statisticiens ou chercheurs. Par ailleurs, une partie non négligeable d'entre eux (entre 7.1 % et 27.6 % entre 1960 et 1985) sont nommés à la tête de centres sous la tutelle du ministère (INRP, CIEP etc.), ce qui leur permet à la fois de se tenir informés à distance de l'évolution de la recherche en éducation et de se positionner au sommet hiérarchique des premières instances centralisant la recherche en sciences de l'éducation en France.

Du fait de la congruence à partir du milieu des années 1970 de plusieurs facteurs institutionnels et cognitifs, l'établissement devient progressivement dans les années 1980 l'unité d'analyse fondamentale du système éducatif (Poupeau, 2003). Tous les évaluateurs investissent ainsi la problématique du fonctionnement des établissements et, dans certains cas, de ses effets sur les résultats des élèves. Certains insistent sur des problématiques précises (comme les chercheurs de l'Iredu, les statisticiens de la DEPP ou les magistrats de la Cour), d'autres entendent étudier plusieurs, voire tous les aspects de ce fonctionnement (experts de l'OCDE, sociologues de l'INRP, et inspecteurs généraux). Dans cette concurrence cognitive, ces derniers proposent jusqu'en 1992 la science de gouvernement la plus générale.

Plusieurs facteurs contribuent à cette domination qui renvoient aussi bien aux inspirations et à la méthodologie des inspecteurs généraux qu'aux interrogations ou initiatives inabouties de la part d'autres évaluateurs. D'un côté, les inspecteurs généraux réfléchissent, non pas de manière continue mais par à coups, à une méthodologie générale d'évaluation des établissements depuis les premières initiatives d'inspections croisées lancées par le doyen de l'Igen<sup>13</sup> en 1976. L'expérience des « inspections paysages » lors du moratoire des inspections individuelles de 1982 et les premières réflexions internationales sur les pratiques d'évaluation des établissements suscitées par le CERI<sup>14</sup> (OCDE) dans les années 1980 dans le cadre du projet ISIP<sup>15</sup> constituent autant de sources d'inspiration pour les inspecteurs des deux inspections générales (Igen et IGAENR<sup>16</sup>) désireux de travailler ensemble. Ce mouvement aboutit en 1989, 1990 et 1991 par trois vagues d'évaluation de 100 établissements commanditées par le ministre. La méthodologie développée à cette occasion permet d'intégrer les travaux produits par d'autres acteurs, notamment les statisticiens du ministère, comme une étape au mieux initiale, au pire inutile (Pons, 2008). De l'autre, aucun acteur n'est véritablement capable de proposer une procédure d'évaluation des établissements formalisée. Les statisticiens du ministère par exemple ne sont pas en mesure de produire des évaluations des acquis des élèves de manière suffisamment régulière pour les utiliser dans une évaluation des effets de l'organisation des établissements sur les résultats des élèves.

Toutefois, cette dynamique s'arrête à la rentrée de l'année 1992. Les éléments qui avaient fait le succès des inspecteurs généraux disparaissent progressivement : baisse du nombre et de la proportion d'inspecteurs généraux exerçant une fonction de direction d'un établissement public, non-reconduction par les décideurs des vagues d'évaluation des établissements et surtout plus forte capacité des autres groupes professionnels à proposer une science de l'évaluation dans un contexte où l'unité d'analyse se déplace de l'établissement vers le système éducatif dans son ensemble. Dans ce processus, les statisticiens du ministère, qui perfectionnent leur *science du constat chiffré* depuis 1987, deviennent progressivement les évaluateurs dominants des années 1990. Depuis la création de la DEP en février 1987 et tout au long de l'ère Thélot (1990-1997), les statisticiens du ministère parviennent en effet dans leurs travaux à intégrer plusieurs savoir et savoir-faire d'autres groupes professionnels, y compris ceux des acteurs les plus dynamiques de la période (comme les chercheurs de l'Iredu par exemple), et ce par différents vecteurs institutionnels : rapatriement des évaluations internationales, développement des commandes à l'intention des chercheurs, partage puis importation de leurs compétences dans le cadre de grands colloques (sur l'investissement en éducation ou la violence) ou séminaires méthodologiques et coordination de multiples comités de pilotage de leurs propres dispositifs d'enquête (évaluations standardisées, indicateurs IPES-IVAL et autres évaluations ponctuelles). En parallèle, certains évaluateurs se replient sur leurs premiers

---

13 Inspection Générale de l'Éducation Nationale.

14 Centre for Educational Research and Innovation.

15 International School Improvement Project.

16 Inspection Générale de l'Administration de l'Éducation Nationale et de la Recherche.

savoir-faire (cas des inspecteurs de l'IGEN notamment) ou se spécialisent dans des domaines connexes (cas des chercheurs du Céreq). À l'inverse, les statisticiens du ministère, par la mise en œuvre de plusieurs publications et outils d'évaluation, sont capables de proposer une vision d'ensemble du système éducatif, dans laquelle les différentes unités d'analyse s'emboîtent (les élèves, les établissements, les académies puis le système dans son ensemble, voire le système éducatif français au regard des autres pays), en tout cas d'un point de vue logique. Enfin, ils sont les seuls à proposer un modèle d'évaluation explicite (théorie de l'effet miroir de Claude Thélot) qui permet de rationaliser les différentes pratiques évaluatives existantes et de les définir de fait comme autant de formes particulières de ce que fait la DEP à grande échelle.

Cette tendance change radicalement à partir de 1997. Après le contexte politique ambivalent caractéristique de la période Allègre (1997-2000), qui a pour conséquence de discréditer les évaluateurs officiels sans pour autant donner plus de pouvoir à d'autres évaluateurs (les chercheurs notamment), tous les groupes connaissent d'importants questionnements méthodologiques, qui prennent trois formes principales. Les anciens groupes dominants tentent d'améliorer la portée générale de leurs travaux, soit en envisageant l'évaluation à une plus grande échelle territoriale (cas des inspecteurs généraux et de l'évaluation de l'enseignement dans les académies) soit en mettant en œuvre plusieurs actions qui ont toutes en commun de tenter de renouer avec un passé glorieux dont on veut faire la synthèse (cas des statisticiens du ministère). Les chercheurs (du Céreq, de l'INRP et de l'Iredu) se spécialisent dans des problématiques qui ont pour conséquence de les écarter progressivement du champ de l'évaluation du système éducatif français pour les emmener soit vers une recherche plus académique soit vers l'évaluation d'autres pays. Enfin de nouveaux groupes parviennent mieux qu'auparavant à proposer leur science de l'évaluation (magistrats de la Cour et experts de l'OCDE).

L'année 2006 constitue un dernier changement important pour les évaluateurs, car la convergence de l'entrée en vigueur de la LOLF le 1er janvier, de la diffusion dans le débat public des impératifs liés à la stratégie de Lisbonne et des interrogations sur les mauvais résultats des élèves français à la troisième évaluation du programme PISA contribue à faire émerger, de façon relativement soudaine, la problématique de l'évaluation de la performance de l'action de l'État, problématique à laquelle les évaluateurs en présence sont inégalement sensibilisés. La mise en œuvre de ces nouveaux instruments d'action publique (loi organique et comparaison internationale), qui viennent se superposer à l'évaluation du système éducatif, implique à la fois un changement d'unité d'analyse et l'apparition de nouveaux évaluateurs, comme les experts de l'OCDE et les inspecteurs généraux des Finances, promouvant de nouvelles approches.

### ***Conclusion : quelle science politique face aux sciences d'État ?***

Même si les deux sciences d'État sur lesquelles nous avons porté notre analyse sont aujourd'hui concurrencées par d'autres sciences de gouvernement, il n'en demeure pas moins qu'elles constituent deux des principales sources de connaissance (sinon les principales sources) disponibles à ce jour quand il s'agit d'analyser le fonctionnement du système éducatif français dans son ensemble. Le sociologue des politiques éducatives ne peut rester insensible à ce constat dans la mesure où l'analyse des politiques publiques propose des outils conceptuels et des cadres théoriques capables d'analyser, souvent de façon macroscopique, les modalités de conception et de mise en œuvre des réformes et les transformations d'ensemble des modes de régulation de l'action publique.

Toutefois, deux écueils le guettent dans son rapport à ce type de connaissance. Dans la mesure où ces sciences d'État reposent sur une légitimité dans certains cas institutionnalisée de longue date, la tentation est forte pour le sociologue de l'action publique de vouloir dialoguer avec elles et répondre, par d'autres approches, à leurs questionnements plutôt que d'envisager ses propres problématiques, et ce d'autant plus quand il existe encore peu de travaux développés dans la même perspective d'analyse que la sienne. Le sociologue des politiques éducatives deviendrait alors une nouvelle figure de l'expert en éducation. À l'inverse, tomber dans un académisme radical au motif que les modes d'administration de la

preuve différent des méthodes classiques en sciences sociales l'amènerait sans doute à ignorer des sources d'information parfois essentielles. La compréhension mutuelle d'approches différentes suffira-t-elle à dépasser cette fausse alternative ?

### **Références**

- A. D. Abbott (1988), *The system of professions : an essay on the division of expert labor*, Chicago, Chicago University Press.
- CURAPP (1993), *L'évaluation dans l'administration*, PUF.
- DEP (2004), *Le pilotage du système éducatif. Enjeux, outils, perspectives*, Paris, MEN, DEP.
- P. Duran, E. Monnier (1992), « Le développement de l'évaluation en France : nécessités techniques et exigences politiques », *Revue française de science politique*, 42, 2, p. 235-262.
- Y. Duterqc (2000), *Politiques éducatives et évaluations : querelles de territoires*, Paris, PUF.
- C. Forestier, C. Thélot, J-C. Emin (2007), *Que vaut l'enseignement en France ?*, Paris, Stock.
- O. Ihl, M. Kaluszynski, G. Pollet (2003), *Les sciences de gouvernement*, Paris, Economica.
- M-C. Kessler, P. Lascoumes, J-C. Thoenig (Dir.) (1998), *Évaluation des politiques publiques*, Paris, L'Harmattan.
- J. Leca (1993), « Sur le rôle de la connaissance dans la modernisation de l'État et le statut de l'évaluation », *Revue française d'administration publique*, n° 66, avril-juin 1993, p. 185-196.
- E. Monnier (1987), *Évaluation de l'action des pouvoirs publics*, Economica.
- J-P. Nioche, R. Poinard (1984), *L'évaluation des politiques publiques en France*, Economica.
- C. Pair (2001), *Forces et faiblesses de l'évaluation du système éducatif en France*, rapport pour le HCEE, octobre 2001.
- A-M. Paveau (1999), « Formes et fonctions de la *doxa* dans les discours sur l'école », *Mots*, n° 61, p. 9-27.
- B. Perret (2003), *L'évaluation des politiques publiques*, Paris, La Découverte.
- X. Pons (2006), « Rôle et usages de l'évaluation en matière de décentralisation », in A. van Zanten (Dir.), *La décentralisation éducative en France*, Paris, OSC.
- X. Pons (2008), *L'évaluation des politiques éducatives et ses professionnels. Les discours et les méthodes (1958-2008)*, thèse de doctorat de science politique, Paris, IEP.
- X. Pons (2009), « Les statisticiens du ministère de l'Éducation nationale : la science depuis l'État ? (1957-2007) », in J-N. Luc, P. Savoie (Dir.), *L'État et l'éducation 1808-2008*, Paris, INRP, à paraître.
- X. Pons, A. van Zanten (Coord.) (2008), *Mapping the knowledge producers on the French education sector. An overview of the evaluation and expertise of the education system*, rapport de recherche pour le projet KnowandPol, juin 2008.
- F. Poupeau (2003), *Une sociologie d'État. L'école et ses experts en France*, Paris, Raisons d'agir.
- G. Simmel (1992), *Le conflit*, Paris, Circé.
- J. Vogler (1996), *L'évaluation*, Paris, Hachette.