

## LA RECHERCHE EST-ELLE SOLUBLE DANS LE CHIFFRE ?

### Contribution au débat sur l'évaluation à la Fondation nationale des sciences politiques

Les sessions de la Commission des chercheurs de la Fondation nationale des sciences politiques qui se sont déroulées les 17 et 30 juin 2009 ont suscité au sein de la communauté scientifique de la FNSP des interrogations et des discussions portant notamment sur les modalités de fonctionnement de la Commission ainsi que sur les critères qui ont été employés pour évaluer l'activité et la production des chercheurs. Il semble que la Commission ait convenu de tenter de tirer les leçons de la première vague d'évaluation conduite selon la nouvelle formule et des réactions qu'elle a provoquées lors d'une prochaine réunion<sup>1</sup>. Je souhaiterais apporter ma contribution à cette réflexion. Je suis, sauf erreur de ma part, le plus ancien<sup>2</sup> chercheur de statut FNSP en activité<sup>3</sup>. Cela ne me confère, évidemment, aucun droit ni privilège. En revanche, l'expérience acquise au fil des ans, la mémoire de l'institution que je conserve et l'attachement<sup>4</sup> que je lui porte m'autorisent peut-être à intervenir dans le débat en cours.

#### L'idéologie de l'évaluation

Ce débat n'est pas propre à la FNSP, il s'insère dans les discussions de grande ampleur qui traversent l'ensemble de la recherche scientifique française à la suite des réformes que des gouvernements successifs ont voulu y introduire depuis plusieurs années, discussions qui se sont matérialisées dans de nombreuses publications<sup>5</sup>. La plupart des opinions qui s'y sont exprimées sont extrêmement critiques vis-à-vis de nombreux aspects de ces réformes et, notamment, des modifications des procédures d'évaluation des chercheurs, de leur activité et de leur production. L'évaluation, en tant que telle, n'y est pratiquement jamais remise en cause mais seulement l'idéologie de l'évaluation inspirée du *New Public Management* qui a transposé des

---

<sup>1</sup> Intervention de Michel Gardette au Conseil d'unité du CERI, le 16 novembre 2009.

<sup>2</sup> Mais non le plus âgé.

<sup>3</sup> J'ai été engagé à la Fondation en octobre 1968, pour collaborer au Centre de recherches administratives dirigé par Jean-Luc Bodiguel et Marie-Christine Kessler, avant d'intégrer le CERI en avril 1969, puis de rejoindre le CEAN en avril 2008. Depuis plus de quarante ans, j'ai donc travaillé sans interruption (sauf de brefs séjours à l'étranger pour enseigner ou diriger un Institut français) à la FNSP.

<sup>4</sup> Je sais ce que je dois à cette institution où j'ai franchi tous les grades, depuis celui d'« assistant de recherche » qui existait autrefois jusqu'à celui de directeur de recherches, qui m'a permis de rédiger deux thèses (3<sup>ème</sup> cycle et État), qui m'a laissé une grande liberté dans la détermination de mes orientations de recherche tout en me fournissant un environnement intellectuel particulièrement encourageant et stimulant, sous la houlette bienveillante des directeurs successifs du CERI et de l'actuel directeur du CEAN.

<sup>5</sup> En ce qui concerne la politologie, voir par exemple : « Recherche et enseignement supérieur, débats et controverses autour de l'évaluation », [www.afsp.msh-paris.fr/observatoire/etudes/evaluation/evaluation.html](http://www.afsp.msh-paris.fr/observatoire/etudes/evaluation/evaluation.html)

normes venues du secteur privé au public, normes qui ont clairement informé les nouvelles procédures d'évaluation appliquées à la recherche scientifique (Aflalo 2009 ; Blay 2009 : 15-16). Deux aspects de ces nouvelles procédures font l'objet de critiques particulièrement appuyées : la fragmentation de l'activité de recherche (Barbériis 2009 : 11) et la notation chiffrée des secteurs ainsi isolés les uns des autres. La fragmentation des activités interdit en effet de saisir dans sa globalité le travail de recherche, donc de prendre en compte la dynamique des interactions qui se nouent entre ses diverses composantes et manifestations (production de « résultats », sous forme de textes ou autres ; enseignement ; diffusion ; animation ; direction et gestion). Le chiffrage équivaut à substituer une appréciation de la quantité à celle de la qualité (Aflalo 2009 : 82-84) et fait que les nouvelles procédures d'évaluation aboutissent « [...] à réduire la diversité et la complexité des activités de recherche à des catégories simples qui évitent de lire les travaux, de consulter les spécialistes, de discuter les critères. » (Charle : 163)

S'agissant de la FNSP, on aurait pu croire que l' « Accord sur la classification et l'évaluation des chercheurs, avenant à l'accord sur les classifications, rémunérations et carrières du 29 mars 2000 » signé le 19 décembre 2008 par l'Administrateur de la Fondation nationale des sciences politiques et les représentants des syndicats mettait les chercheurs de cette institution à l'abri des implications les plus dangereuses de cette idéologie de l'évaluation. Cet accord prévoyait en effet que « Le rapport d'activité est un recensement précis et actualisé de la production scientifique et des activités de chaque chercheur, ils permettent [sic] d'éclairer la Commission lorsque celle-ci doit se prononcer sur les carrières des chercheurs (notamment pour les augmentations au mérite et les promotions) mais ces documents ne sauraient être réduits à cette seule finalité. La fonction première et essentielle des rapports d'activité est de nature scientifique dans la mesure où ils apportent des éléments d'information précis et actualisés sur la production et sur les activités scientifiques des chercheurs. » (p. 4) Or on est en droit de se demander si le traitement quantitatif appliqué par la Commission des chercheurs en juin 2009 aux rapports d'activité a été véritablement « éclairant » et a permis de satisfaire à la « fonction première », de « nature scientifique » que leur conférait l'Accord.

## **L'évaluation à la FNSP en juin 2009**

La manière dont a procédé la Commission en juin 2009 soulève un certain nombre de questions importantes<sup>6</sup>.

La première concerne l'application *a posteriori* de critères qui n'ont pas été publiés au début de la période d'évaluation. La Commission a en effet examiné des rapports décrivant une activité sur deux ans en leur appliquant des grilles d'analyse adoptées à la fin de la période évaluée, grilles qui, au surplus, ont semble-t-il évolué entre la première session (17 juin) et la seconde (30 juin). Dans ces conditions

---

<sup>6</sup> Il ne s'agit absolument pas ici de mettre en cause personnellement les membres de la Commission des chercheurs mais d'interroger le fonctionnement de la Commission en tant que telle et, notamment, les critères qui ont été retenus pour procéder aux évaluations.

(même si elles étaient les mêmes pour tous les chercheurs évalués), l'adéquation des critères aux activités, et aux rapports rédigés pour en rendre compte, ne pouvait être qu'imparfaite. D'autant plus que ces critères n'ont jamais été véritablement précisés, au delà du plan général selon lequel les rapports devaient être construits. On peut en voir un exemple dans l'incertitude qui règne quant à l'importance attribuée aux supports de publication. L'Accord du 19 décembre 2008 en fournit deux listes non numérotées mais dont on peut supposer qu'elles sont implicitement classées par ordre de valeur, la première dans le corps de l'Accord, la seconde dans l'annexe 2 concernant les rapports d'activité. La première évoque :

« le nombre et la qualité des publications scientifiques, en langue française, anglaise ou autres, en prenant en compte :

- . les ouvrages de recherche en nom propre ;
- . les articles dans des revues à comité de lecture ;
- . la coordination d'ouvrages ;
- . la coordination de numéros spéciaux de revue ;
- . les contributions à des ouvrages collectifs. » (Article 3-2-1-2, p. 6)

La seconde demande au « salarié » de présenter :

« la liste de ses productions scientifiques en distinguant :

- ouvrage(s) de recherche ;
- direction(s) d'ouvrage ;
- article(s) de recherche dans des revues avec comité de lecture ;
- contribution(s) à des ouvrages collectifs de recherche ;
- article(s) dans des revues sans comité de lecture ;
- communication(s) (publiées) dans des congrès, colloques, symposiums ;
- autres. » (Annexe 2, rapports d'activité, pp. 9-10)

On constate que les places des directions d'ouvrages et des contributions à des ouvrages collectifs sont différentes dans ces deux passages.

Les grilles d'analyse utilisant des critères flous ont servi à examiner 58 rapports en deux sessions, dont la première a passé beaucoup de temps à mettre au point les modalités de travail de la Commission. Les membres de la Commission sont évidemment arrivés en séance en ayant lu attentivement l'ensemble des rapports mais, même en tenant compte de cet examen préalable, il paraît difficile de croire que la production de chacun des chercheurs examinés ait bénéficié du temps nécessaire à une évaluation qualitative approfondie, que l'on est en droit d'attendre lorsque les décisions de la Commission ont des conséquences immédiates sur les rémunérations et, à plus long terme, sur la « réputation » et les promotions de grade. Cet examen des rapports effectué, selon un membre de la Commission, « au pas de course » semble avoir été dénué de la sérénité dans laquelle il aurait dû se dérouler. Il paraît, en tout cas, qu'il n'a pas été accompagné de la lecture des travaux mentionnés et n'a pu que les comptabiliser, retombant ainsi dans les travers de l'idéologie de l'évaluation chiffrée.

Les décisions de la Commission ont été communiquées aux chercheurs sous forme d'une « fiche individuelle d'évaluation », accompagnée d'une lettre

impersonnelle<sup>7</sup> et technique signée de l'Administrateur de la FNSP. La brutalité de cette forme de communication des résultats peut avoir (et paraît avoir eu dans quelques cas) des effets néfastes sur le moral et la confiance en soi de celles et ceux qui ont appris de cette manière qu'ils avaient été « mal » évalués. Cela ne semble pas constituer l'incitation à « améliorer » ou « intensifier » le travail que l'idéologie de l'évaluation inspirée du *New Public Management* prétend prôner. Plus généralement, si quelque 56 % des chercheurs ont fait l'objet d'une évaluation positive entraînant une augmentation de salaire, cela signifie que environ 44 % des chercheurs ont, eux, été jugés « insuffisants ». Que presque la moitié des chercheurs d'une institution telle que la FNSP soient jugés tels devrait inquiéter. De fait, l'évaluation est en général conçue comme une procédure qui doit avoir pour but d'aider et non de sanctionner les personnes évaluées ; d'entraîner et d'inciter plutôt que de décourager et de démoraliser (Champagne 2003 : 3-4). La Directrice des relations humaines du CNRS explique que, désormais, la filière RH est chargée de « piloter » un « processus de suivi » des chercheurs ayant fait l'objet d'une « avis réservé » ou d'une « alerte » et que ce processus a donné des résultats satisfaisants (D'Argouges 2009). À la FNSP, l'Accord du 19 décembre 2009 prévoit que : « [...] lorsqu'un rapport apparaît aux membres de la Commission comme étant nettement insuffisant (dans la forme comme sur le fond) celle-ci peut demander aux directeurs des centres de recherche concernés de prendre toutes les dispositions nécessaires pour remédier à ces **insuffisances**. » (p. 4, en gras dans le texte) À ma connaissance, cela n'a pas été le cas et les quelques demandes d'explication adressées par des chercheurs aux autorités de la FNSP quant à l'évaluation dont ils ont été l'objet n'ont pas reçu de réponse.

## L'évaluation du mérite

Dans la terminologie adoptée par la FNSP pour définir les fonctions de la Commission des chercheurs, celle-ci formule des recommandations quant aux « augmentations au mérite ». Or une étude de l'emploi de la notion de « mérite » dans les administrations publiques (Ihl 2010) montre combien elle est floue et malléable et le philosophe Yves Michaud (Michaud 2009) a minutieusement mis au jour les mécanismes d'une évolution qui l'a, aujourd'hui, ou vidée de son sens, ou pervertie : « Notre idée du mérite a été, pour ainsi dire, filtrée et appauvrie : de capacités récompensées par des attributions de responsabilité et d'actions bonnes récompensées par l'estime (des autres et de soi), on est passé à des mérites individuels dont on attend le salaire [...] le mérite opère aujourd'hui comme une sorte d'ornement ou de décoration destinée à donner un semblant de justification à des classements et des hiérarchies qui n'ont rien de méritoire. Nous en appelons au mérite parce que nous voulons faire des différences entre les individus sans parvenir à les justifier faute à la fois d'un cadre stable d'évaluation et de critères déterminés d'évaluation. » (pp. 6 et 239) L'identification et la comptabilité des mérites se révèle toujours fragile et approximative ; elle est « condamnée à être inexacte » (p. 270) et ne saurait être basée sur des indicateurs chiffrés. Car, si le mérite demeure « une

---

<sup>7</sup> Au point qu'une chercheuse dont le prénom ne peut laisser place à la moindre équivoque s'est vue adresser comme « Monsieur »...

catégorie indispensable pour notre appréhension des situations humaines en tant qu'humaines » (p. 269), selon François Bourricaud (1990), elle n'a de sens que pour « un petit nombre de professions à responsabilité forte » (cité dans Michaud 2009 : 40) et les métiers de la recherche ne répondent pas à cette description.

Ce que mesurent les critères adoptés par la Commission des chercheurs ne peut donc, dans cette conception, être le mérite. La « fiche individuelle d'évaluation » consiste en une suite de notes attribuées à l'activité ou à la production des chercheurs selon un découpage particulier de leur travail. Le coefficient le plus important y est accordé aux publications. Je ne vais pas reprendre ici les critiques montrant que, derrière les indicateurs bibliométriques et le calcul des « facteurs d'impact », se cachent en réalité des procédures discordantes, incertaines (Briatte 2008 ; Büttgen, Cassin 2009 ; Debrégeas, Jobard 2008), manipulables (Kermarrec et al. 2007), dangereuses (Cingras 2008), dont la généralisation contient un risque grave d'incitation au conformisme (Audier 2009 ; Blay 2009 : 23-24 ; Charle 2009 : 166-167). Ceci est maintenant bien connu et les exemples ne manquent pas d'articles cités de nombreuses fois en illustration d'erreur à ne pas commettre qui, de ce fait, ont un facteur d'impact élevé ou de résultats erronés non détectés par les comités de lecture des revues scientifiques les plus prestigieuses<sup>8</sup>.

La Commission des chercheurs paraît avoir entendu, au moins en partie, ces critiques puisqu'elle n'a pas établi de liste hiérarchisée de publications. De ce fait, les critères utilisés pour aboutir à la notation chiffrée de la case « publications » demeurent flous et ne peuvent permettre aux chercheurs de comprendre comment ils ont été notés. Le classement par ordre d'importance des supports de publication étant, on l'a vu, incertain, les chercheurs demeurent dans le brouillard quant à ce qu'ils auraient dû faire, à ce qu'ils devraient faire pour être « bien » notés.

Les mêmes incertitudes flottent sur la notation dans les autres cases. Est-ce le nombre des enseignements qui est évalué ? Leur qualité ? Sur quelle base : les fiches d'évaluation remplies par les étudiants ? Mais alors, qu'en est-il de ceux qui enseignent en dehors de l'Institut d'études politiques de Paris ? Comment juge-t-on de l'insertion dans les « réseaux » ? À partir de la participation à des groupes de travaux formalisés ? De la participation aux colloques ? Aux ouvrages collectifs ? Faudra-t-il que chaque chercheur annexe à son rapport d'activité son carnet d'adresses professionnelles ? Dans le domaine des « réseaux » il n'y a clairement pas adéquation du critère à son objet<sup>9</sup>. La case « administration et animation de la recherche » associe deux activités extrêmement différentes, même si elles se recoupent parfois ; ici l'indicateur manque complètement de cohérence<sup>10</sup>. Comment apprécier la valorisation ? Enfin, que faut-il entendre par prix ? Heureusement, cet inventaire n'inclut pas les brevets, comme le faisaient encore récemment certaines fiches du CNRS !

---

<sup>8</sup> Voir, par exemple, sans remonter aux polémiques déclenchées par Alan Sokal : « L'immense supercherie d'un chercheur en cristallographie, l'Université d'Alabama met en cause les "découvertes" de l'un de ses biologistes moléculaires », *Le Monde* 26 décembre 2009 : 13.

<sup>9</sup> « Une première propriété d'un bon indicateur est son adéquation à l'objet. » (Cingras 2008 : 6).

<sup>10</sup> « Une seconde propriété importante qu'un bon indicateur doit posséder est d'être homogène dans sa composition. » (Cingras 2008 : 7)

Au total, qu'il s'agisse des publications ou des autres secteurs d'activité des chercheurs, aucune base objective n'existe pour formuler des notes fragmentées dont l'agrégation aboutit à des résultats *injustifiables*, au sens propre du terme. Il ne sera pas possible de sortir d'une situation dans laquelle le flou et l'incertitude pourraient conduire à la suspicion d'arbitraire si la sectorisation des activités de recherche est maintenue comme base de l'évaluation des chercheurs. Et les leçons que l'on commence à tirer des méfaits de l'évaluation chiffrée dans l'administration publique devraient inviter à réfléchir sur ses conséquences pour la recherche : « De nombreux travaux académiques témoignent déjà de manière rigoureuse des risques de ces méthodes<sup>11</sup> (en terme d'impact sur la performance — risque de focalisation sur le court terme, impact négatif sur les capacités d'innovation, la coopération ou la coordination, la confusion entre l'évaluation des activités et des individus — et d'effets sur les individus — la corrélation entre ces pratiques et l'accroissement de la souffrance au travail est de plus en plus documentée, en particulier par la psychopathologie du travail) et de l'existence de nombreux autres moyens de pilotage de la performance (par les règles, par les pairs, etc.) souvent déjà largement mis en œuvre avec succès dans les organisations dans lesquelles on souhaite introduire la culture du résultat. » (Dobler, Saulpic 2009)

## **Quelques propositions**

L'Accord du 19 décembre 2008 contient une clause de révision (article 5) ; en outre, la pratique effective de la Commission des chercheurs demeure en grande partie en dehors du champ couvert par cet accord, notamment en ce qui concerne la sectorisation des activités et leur appréciation chiffrée. Une modification de cette pratique est par conséquent aisément envisageable. Dans cette perspective, je souhaiterais présenter quelques propositions.

### **La science comme travail**

Il faut sans doute repartir de l'idée, défendue par le spécialiste de la psychologie du travail qu'est Christophe Dejourn<sup>12</sup> (2003), que la science doit être prise comme un travail et qu'elle doit être considérée comme telle dans l'évaluation : « La science est elle aussi un travail. L'évaluation de l'activité scientifique est loin d'être satisfaisante. D'une part, elle ne sanctionne que les résultats de la recherche et non le travail, les échecs, la souffrance, la mobilisation subjective et l'opiniâtreté qu'elle implique. D'autre part, par les critères d'évaluation, elle génère des effets pervers. La course effrénée aux publications et aux communications scientifiques et surtout la multiplication des revues, des supports spécialisés et des colloques confine à l'absurde. » (pp. 54-55) Il faudrait probablement aussi retrouver le sens classique du « mérite » que rappelle Yves Michaud (2009) : « comme accomplissement d'une

---

<sup>11</sup> Pratiques des évaluations des individus sur la base d'un nombre le plus réduit possible d'indicateurs chiffrés, associées à des systèmes formels de sanction-récompense.

<sup>12</sup> Directeur du Laboratoire de psychologie du travail et de l'action et professeur au Conservatoire national des arts et métiers (CNAM).

vie plutôt que comme comptabilité des actions. » (p. 247) La considération du mérite ne pourrait donc valoir qu'au moment du passage à l'*éméritat*, voire peut-être à la « classe exceptionnelle », mais ne saurait intervenir dans des décisions d'augmentation de salaire résultant de l'examen d'une activité limitée dans le temps.

### **Des périodes d'activité « contextualisées »**

On a débattu de la période la plus pertinente pour l'évaluation des activités de recherche, certains pensent que quatre années permettraient une appréciation plus sérieuse que deux ans. Dans tous les cas, évaluation des activités sur une durée de deux ou de quatre années, il est impossible d'ignorer complètement le « contexte historique » de ces activités, c'est-à-dire comment se situe cette durée dans un cycle plus long d'organisation du travail qui peut inclure des périodes de recherches sur le terrain (en mission ou en détachement), des périodes de réflexion, de réorientation théorique, d'essais méthodologiques, de familiarisation avec un nouveau terrain et de collecte d'informations<sup>13</sup> ; ne pas publier pendant deux ou quatre ans n'est pas nécessairement signe d'activité insuffisante et ne devrait pas pénaliser les chercheurs qui, à un moment de leur carrière, doivent laisser mûrir un travail. Comme nous le savons tous, la recherche a besoin de liberté (Blay 2009 : 19) et de longue durée ; séquencer l'évaluation en périodes de deux, ou de quatre, ans n'est acceptable que si ces périodes sont comprises (dans tous les sens du terme) dans une dynamique de recherche au long cours.

### **Une évaluation globale sans recours à des indicateurs chiffrés**

Cette évaluation « contextualisée » devrait être globale et non fragmentée ; elle devrait abandonner l'utilisation d'indicateurs chiffrés. Dans le secteur public, « [...] de la même manière qu'il peut encourager certains comportements contre-productifs au lieu de les limiter, un tel pilotage par le chiffre peut conduire tout aussi ironiquement à ce que le *monitoring* de l'activité devienne sa norme unique, au détriment de la poursuite des objectifs en eux-mêmes. » (Guillaume 2009 :108) Dans la recherche, « [...] si les indicateurs peuvent donner des tendances sur un nombre réduit d'aspects de la vie scientifique, il convient d'être très circonspect dans leur usage en raison de la possibilité d'interprétations erronées, des erreurs de mesure (souvent considérables) et des biais dont ils sont affectés. Un usage abusif des indicateurs est facilité par la nature chiffrée du résultat qui introduit la possibilité d'établir dans l'urgence toutes sortes de statistiques, sans se préoccuper d'en analyser la qualité et le contenu, et en occultant l'examen d'autres éléments de la vie scientifique comme, par exemple, l'innovation et le transfert intellectuel et industriel. » (Kermarrec et al. 2007 : 5) Cela est particulièrement vrai pour les sciences humaines et sociales. L'utilisation d'indicateurs chiffrés est porteuse de risques que l'on connaît dans les entreprises privées, dans le secteur public et que l'on voit se profiler dans la recherche : « Le risque inhérent [aux nouvelles modalités

---

<sup>13</sup> Il conviendrait, en outre, de prendre en compte la nature des domaines et terrains de recherche ; travailler sur des pays étrangers, parfois lointains, implique une familiarisation, des apprentissages qui demandent plus de temps que de travailler sans sa société d'origine et dans sa langue natale. Et, sans parler de « prime de risque », les dangers encourus sur certains terrains devraient aussi entrer dans l'évaluation.

d'évaluation] est de voir se mettre en place des stratégies de contournement ou de détournement dont on sait les résultats délétères<sup>14</sup> : sur-publication (publier plusieurs fois des résultats quasi-identiques), auto-citation ou citation entre collègues, publication précoce de résultats non stabilisés, auto-génération de bruits de fond artificiels (et négligence des pistes silencieuses mais fructueuses), court-circuitage des lieux d'évaluation non administratifs (les colloques ou séminaires), etc. Tout privilégie ici une science bling-bling, une science caisse-enregistreuse, une science pointeuse, où les bips de citation se substituent à l'examen des avancées et des promesses réelles de l'activité. » (Debrégeas, Jobard 2008)

Il convient d'échapper autant que faire se peut aux effets « Prix Goncourt » (stratégie d'adaptation aux critères d'évaluation) et « acrobate de cirque » (dépenser autant d'énergie à faire savoir la difficulté de ce qu'on fait qu'à faire réellement), paradigme proposé par Erwin Goffman (Champagne 2003 : 7), pour privilégier ce qui doit être le but final de la recherche : le projet, au sens que lui donnait Gaston Bachelard (1963 : 11), l'innovation et les débats qu'elle suscite. Comme le suggérait ce dernier, « [...] l'esprit scientifique est essentiellement une rectification du savoir, un élargissement des cadres de la connaissance. Il juge son passé historique en le condamnant. Sa structure est la conscience de ses fautes historiques. Scientifiquement, on pense le vrai comme rectification historique d'une longue erreur, on pense l'expérience comme rectification de l'illusion commune et première. » (Bachelard 1963 : 173) Sur cet arrière plan, le sens que donne Christophe Dejours (2003 : 15 et 16) à travailler et à « bien » travailler (s'écarter des prescriptions, faire des infractions) rejoint une conception de l'excellence qui a peu à voir avec celle des « pôles » autour desquels on prétend à l'avenir organiser la recherche : « [...] nous voulons, disent Barbara Cassin et Philippe Büttgen, que les meilleurs émergent. Mais nous voulons qu'ils émergent comme ils sont, c'est-à-dire bizarres et non pas conformes ou déjà conformés (certifiés conformes) par les *rankings* de revues ou d'universités. » (Büttgen, Cassin 2009 : 39) Remettre en question les héritages, faire advenir le nouveau comme imprévu et éventuellement « bizarre », c'est effectivement en cela que devrait consister le *travail* de recherche pour conserver à ce mot sa dimension de production de connaissance. C'est précisément ce qui ne peut être apprécié à l'aide d'indicateurs chiffrés fragmentaires. Toute évaluation, qu'elle porte sur une période d'activité limitée ou sur l'ensemble d'une carrière, devrait être inspirée par cet esprit scientifique.

### **Basée sur la lecture, la connaissance des activités et la délibération**

Pratiquement, il conviendrait d'abandonner la notation sectorielle pour en revenir à une appréciation globale de l'activité des chercheurs. Les indicateurs chiffrés ne sont pas moins flous, sans doute même le sont-ils davantage, qu'une évaluation d'ensemble issue d'une discussion dans laquelle interviennent des collègues qui ont une connaissance de première main de l'activité du chercheur au sein d'un laboratoire. Les publications doivent tenir une place importante dans cette

---

<sup>14</sup> Pour une critique de ces effets au-delà de la seule recherche, voir « Le gouvernement des conduites, l'expérience blairiste du pouvoir », entretien avec Patrick Le Galès et Florence Faucher-King, *Vacarme* 38, hiver 2007 : 57-61.

évaluation mais, pour pouvoir la tenir, elles doivent impérativement être lues ; on ne saurait en aucun cas se contenter d'une comptabilité bibliométrique, plus ou moins bricolée en l'absence de liste hiérarchisée de référence<sup>15</sup> : « [...] l'évaluation scientifique est d'abord analyse des contenus scientifiques par les pairs, et [...] le volume individuel ou collectif de publications scientifiques est un indicateur (parmi d'autres) de l'activité scientifique, mais pas de la qualité de l'activité. » (Jobard 2009) Les supports de publications doivent par conséquent importer moins que le contenu des textes. L'activité collective et transdisciplinaire qui se caractérise, notamment, par la direction ou la participation à des ouvrages collectifs devrait être valorisée plus qu'elle ne l'est actuellement. La publication dans des revues ou des collections situées hors du champ conventionnel de la science politique, relevant d'autres disciplines des sciences sociales, devrait être reconnue comme moyen de diffuser des analyses de politologie hors des limites qui lui sont assignées. Dans nos disciplines, la frontière entre ouvrage de recherche et de vulgarisation est extrêmement mince, et l'on sait des ouvrages supposés de vulgarisation qui fournissent l'occasion de proposer, dans des formes accessibles, des idées nouvelles. Une réflexion particulière devrait porter sur les publications électroniques, non seulement parce que, dans une période de crise de l'édition scientifique, elles offrent des moyens puissants de diffusion de la connaissance mais aussi parce qu'elles permettent la combinaison de plusieurs modes de communication qui enrichit considérablement les savoirs transmis. L'expérience d'une collection telle que « Questions de recherche » au CERI, autorisant le multi-media tout en appliquant les mêmes exigences éditoriales que les revues « à comité de lecture » aux textes proposés devrait être source d'enseignements ; les publications électroniques répondant aux mêmes critères de rigueur que les publications papier ne devraient pas être dévalorisées.

### **Prendre le temps d'évaluer**

Lire, réfléchir à ce qu'on a lu (Marzano 2009 : 63-64), en débattre en intégrant à la discussion des éléments de connaissance fiables sur les activités qui accompagnent les publications, tout cela exige qu'on n'examine plus les rapports d'activités dans l'urgence, « au pas de course »<sup>16</sup>, mais que soit donné aux évaluateurs le temps nécessaire pour évaluer sereinement, temps qui, logiquement, devra être pris en compte lorsque les chercheurs qui siègent à la Commission seront à leur tour évalués<sup>17</sup>. Les résultats des évaluations, et leurs conséquences sur les rémunérations, devraient être communiqués avec des commentaires explicatifs et justificatifs, assortis de suggestions adressées aux chercheurs trouvés

---

<sup>15</sup> J'entends évidemment ici qu'il faudrait abolir toute idée de liste hiérarchisée, et non en constituer une « solide »...

<sup>16</sup> Cela vaut pour la Commission des chercheurs de la FNSP mais aussi, plus largement pour l'ensemble des procédures d'évaluation dans le dispositif de recherche scientifique français, au sein duquel les visites « coup-de-poing » des missions de l'AERES figurent en bonne place. En moyenne, le temps pris pour l'évaluation est sensiblement plus long dans les autres pays ayant un secteur de recherche aussi développé qu'en France (Charle 2009 : 161).

<sup>17</sup> Il paraît envisageable de demander aux chercheurs évalués de sélectionner deux ou trois de leurs travaux publiés au cours de la période de référence, qu'ils jugent particulièrement représentatifs ou importants, et de demander à des rapporteurs choisis au sein de la Commission des chercheurs de lire ces travaux.

« insuffisants » afin de transformer ce qui peut être ressenti comme l'expression d'une dévalorisation méprisante en réévaluation positive de soi conduisant à des interrogations sur les pratiques de travail et, par là, à leur amélioration, d'un point de vue scientifique. L'important est ici que les évaluateurs soient placés dans des conditions où ils aient le temps de jauger posément les conséquences (pas seulement salariales mais morales) de leur jugement, et dans lesquelles ils ne soient pas conduits à décréter trop rapidement le « mérite » ou l'« insuffisance » de leurs collègues mais où ils puissent contribuer, par le débat sur la pratique de leurs collègues, à la construction des savoirs car, dans la recherche comme ailleurs, peut-être plus qu'ailleurs, « le ressort fondamental de la mobilisation dans le travail, c'est l'attente du sujet par rapport à l'accomplissement de soi. » (Dejours 2003 : 52). Ceci ne signifie pas que les augmentations de salaire sont négligeables (d'autant moins que, dans la succession des périodes d'évaluation, elles seront cumulatives) mais qu'elles ne sauraient être dissociées de l'encouragement à « l'accomplissement de soi ». Enfin, à l'occasion de l'examen des candidatures à des changements de grade, il serait probablement judicieux d'entendre les postulants car, en des circonstances aussi importantes, pour avoir « accès à la connaissance du travail », il est nécessaire de « passer par la subjectivité des travailleurs », c'est-à-dire par leur parole (Dejours 2003 : 30).

### **La représentation des laboratoires en région**

Outre ces quelques mesures d'ordre général, il serait souhaitable de repenser la représentation des laboratoires en région. Les centres de recherche en région sont légèrement sous-représentés dans la composition actuelle de la Commission<sup>18</sup> ; au surplus, le système de désignation des membres de la Commission représentant les laboratoires en région fait qu'il se trouve toujours l'un d'entre eux pour n'être pas du tout représenté<sup>19</sup>. Cela signifie que, quels que soient les efforts faits par les élus siégeant à la Commission pour s'informer des activités de leurs collègues d'un autre centre, ceux-ci n'en peuvent avoir une connaissance fine<sup>20</sup>. En conséquence, il serait sans doute judicieux d'adopter un mécanisme plus « fédéral » permettant à tous les laboratoires d'être représentés au sein de la Commission et rééquilibrant le rapport Paris / régions.

\* \* \*

J'ai pris la liberté de présenter ces quelques pistes de réflexion parce qu'il m'a paru, comme à nombre de collègues, que la Commission des chercheurs de la FNSP réunie les 17 et 30 juin 2009 avait procédé dans l'examen des rapports d'activités d'une manière qui s'était avérée contraire aux finalités de la recherche

---

<sup>18</sup> Si l'on compte ensemble le directeur et les élus, les centres parisiens bénéficient d'un représentant pour 4 chercheurs et les centres en région d'un représentant pour 5 chercheurs.

<sup>19</sup> C'est le cas, à l'heure actuelle, du CEAN de l'IEP de Bordeaux.

<sup>20</sup> Les conséquences de ce système seraient encore plus gênantes si le directeur de laboratoire en région et le représentant élu des chercheurs en région siégeant à la Commission émanaient tous deux des laboratoires de l'IEP de Bordeaux, Grenoble, c'est-à-dire le PACTE, n'aurait alors aucun représentant.

scientifique et qui avait ébranlé, au moins chez quelques collègues, les perceptions qu'un chercheur peut avoir de lui-même et de son métier. J'ai pleinement conscience que les pistes que je propose vont dans un sens contraire aux directions dans lesquelles les réformes de la recherche voulues par le gouvernement, mais contestées par des secteurs très larges de la communauté scientifique, ont engagé notre métier. Le statut particulier de la Fondation nationale des sciences politiques lui conserve, de ce point de vue, une certaine marge de manœuvre ; la singularité de cette institution a consisté, notamment, à utiliser ces marges de manœuvre pour construire son prestige, auquel la qualité des recherches qui y ont été, et qui continuent d'y être produites, n'a pas peu contribué. Cette marge de manœuvre peut encore être utilisée aujourd'hui pour proposer un modèle d'évaluation qui soit véritablement évaluatif, incitatif et intellectuellement stimulant, mais cela implique de revenir sur certaines des pratiques qui ont été instaurées depuis juin 2009, et de réfléchir aux possibilités de mettre en place une évaluation globale, non chiffrée et correctement informée du travail accompli par les chercheurs. Je souhaite que ces quelques pages puissent contribuer à enrichir le débat sur l'évaluation à la FNSP.

Denis-Constant MARTIN  
Sciences Po Bordeaux (CEAN)  
Le 4 février 2010

## Textes cités :

AFLALO (Aflalo), « L'évaluation : un nouveau scientisme », *CITÉS, philosophie, politique, histoire*, 37, 2009 : 79-89.

AUDIER (Florence), L'évaluation et les listes de revues, 15 septembre 2009, [www.laviedesidees.fr/l-evaluation-et-les-listes-de.html](http://www.laviedesidees.fr/l-evaluation-et-les-listes-de.html) (téléchargé le 12/01/2010)

BACHELARD (Gaston), *Le nouvel esprit scientifique*, Paris, Presses universitaires de France 1963.

BARBÉRIS (Isabelle), « Présentation, le cauchemar de Paul Otlet », *CITÉS, philosophie, politique, histoire*, 37, 2009 : 9-11.

BLAY (Michel), « L'évaluation par indicateurs dans la vie scientifique : choix politique et fin de la connaissance », *CITÉS, philosophie, politique, histoire*, 37, 2009 : 15-25.

BOURRICAUD (François), « Les antinomies de la méritocratie », *Études* 4, octobre 1990 : 335-344.

BRIATTE (François), « Comparaison inter-classements des revues en sociologie-démographie et en science politique », *Bulletin de méthodologie sociologique* 100, octobre 2008, [www.bms.revues.org/index3473.html](http://www.bms.revues.org/index3473.html) (téléchargé le 12/01/2010)

BÜTTGEN (Philippe), CASSIN (Barbara), « "J'en ai 22 sur 30 au vert", six thèses sur l'évaluation », *CITÉS, philosophie, politique, histoire*, 37, 2009 : 27-41.

CHAMPAGNE (Patrick), « Préface », in : DEJOURS (Christophe), *L'évaluation du travail à l'épreuve du réel, critique des fondements de l'évaluation*, Paris, INRA éditions, 2003 : 3-7.

CHARLE (Christophe), « L'évaluation des enseignants-chercheurs, critiques et propositions », *Vingtième siècle, revue d'histoire*, 102, avril-juin 2009 : 159-170.

CINGRAS (Yves), « Note de recherche : La fièvre de l'évaluation de la recherche, du mauvais usage de faux indicateurs », CIRST, 2008-05, [www.cirst.uqam.ca](http://www.cirst.uqam.ca) (téléchargé le 12/01/2010)

D'ARGOUGES (Christine), « La vision du management des chercheurs se transforme », *aef.info*, dépêche n° 119224, lundi 28 septembre 2009, [www.aef.info/public/fr/abonne/depeche/depeche\\_detail.php?id=](http://www.aef.info/public/fr/abonne/depeche/depeche_detail.php?id=) (téléchargé le 28/09/2009).

DEBRÉGEAS (Georges), JOBARD (Fabien), « Vos papiers ! la science à l'aune de la raison comptable », *Vacarme* 44, été 2008, [www.vacarme.org/article1630.html](http://www.vacarme.org/article1630.html) (téléchargé le 12/01/2010)

DEJOURS (Christophe), *L'évaluation du travail à l'épreuve du réel, critique des fondements de l'évaluation*, Paris, INRA éditions, 2003.

DOBLER (Philippe), SAULPIC (Olivier), « L'échec de l'évaluation des ministres ou les limites de la culture du résultat », *Le Monde* 18 août 2009 : 13.

GUILLAUME (Bertrand), « Indicateurs de performance dans le secteur public : entre illusion et perversité », *CITÉS, philosophie, politique, histoire*, 37, 2009 : 101-109.

IHL (Olivier), « Les républiques du concours, l'identification du mérite bureaucratique en France et aux États-Unis », in : MARTIN (Denis-Constant) dir., *L'identité en jeux, pouvoirs, identifications, mobilisations*, Paris, Karthala (Recherches internationales), à paraître en 2010.

JOBARD (Fabien), *Regards sur la production quantitative des politistes de la section 40 du CNRS (2003-08) enrichis de considérations opportunes sur les revues scientifiques et graphiques en couleurs délicates*, [www.congresafsp2009.fr/modules/jobard.pdf](http://www.congresafsp2009.fr/modules/jobard.pdf) (téléchargé le 12/01/2010)

KERMARREC (A.M.), FAOU (E.), MERLET (J.P.) (rapporteur), ROBERT (P.), SEGOUFIN (L.), *Que mesurent les indicateurs bibliométriques ?*, Document d'analyse de la Commission d'évaluation de l'INRIA, Institut national de recherche en informatique et en automatique, 2007, [www.inria.fr/inria/organigramme/ce\\_indicateurs.fr.html](http://www.inria.fr/inria/organigramme/ce_indicateurs.fr.html) (téléchargé le 12/01/2010)

MARZANO (Michela), « *Publish or perish* », *CITÉS, philosophie, politique, histoire*, 37, 2009 : 59-64.

MICHAUD (Yves), *Qu'est-ce que le mérite ?*, Paris, Bourin éditeur, 2009.