

L'évaluation de la recherche en question

Bernard THOMANN

La question de l'évaluation de la recherche est au cœur du mouvement de protestation qui anime les universités et les laboratoires depuis de longs mois. Plusieurs revues lui consacrent des numéros spéciaux : au-delà du caractère massif des oppositions aux réformes, une lecture attentive révèle des divergences d'appréciation sur le rôle de l'évaluation et sur la relation entre science et société.

Recensés :

- « La fièvre de l'évaluation », numéro spécial de la *Revue d'histoire moderne et contemporaine*, supplément n° 55-4 bis, 2008.
- « L'idéologie de l'évaluation, la grande imposture », numéro spécial de *Cités*, n° 37, 2009.
- Christophe Charle, « L'évaluation des enseignants-chercheurs », *Vingtième Siècle. Revue d'Histoire*, n° 102, avril-juin 2009, p. 159-170.
- Patrick Fridenson, « La multiplication des classements », *Le Mouvement Social*, n° 226, janvier-mars 2009, p. 5-14.

La question de l'évaluation a été placée, depuis quelques mois, sous les feux de l'actualité par le mouvement social qui s'oppose au décret instaurant la modulation du service des enseignants-chercheurs en fonction de leurs « performances » individuelles. Cette question s'était néanmoins déjà imposée depuis plusieurs années au cœur de la pratique quotidienne des universitaires et des chercheurs. Qu'elle se traduise par les initiatives des

étudiants pour évaluer les enseignements, par l'utilisation grandissante des indices bibliométriques, par la généralisation de la pratique des appels d'offre de l'Agence nationale de la recherche (ANR), ou par les visites de l'AERES (Agence d'évaluation de la recherche et de l'enseignement supérieur), la rapidité de l'évolution des pratiques de l'évaluation et la multiplication de ses procédures suscitent des interrogations et des inquiétudes. Celles-ci semblent s'être soudainement et violemment cristallisées à la suite de la présentation du projet de décret de réforme sur le statut des enseignants-chercheurs. La prise à bras le corps de cette question de l'évaluation par des numéros spéciaux de la *Revue d'histoire moderne et contemporaine* sur « la fièvre de l'évaluation », et de la revue *Cités* sur l'« idéologie de l'évaluation : la grande imposture », et par des articles du *Mouvement Social* et de la revue *Vingtième Siècle*, a le mérite de venir éclairer l'actualité la plus brûlante en la replaçant dans le temps plus long de l'évolution des politiques de la recherche et de l'enseignement supérieur. Le traitement approfondi de cette question par des revues de sciences humaines et sociales est d'autant plus bienvenu que les médias de grande diffusion peinent parfois à s'élever au dessus de la communication gouvernementale, des caricatures ou des préjugés, et à véritablement rendre compte des enjeux profonds du débat, comme le compte rendu du numéro spécial de *Cités* paru dans *Le Monde* le 21 mars dernier, avec son titre dédaigneux, « Chercheurs qu'on martyrise », l'illustre malheureusement. La lecture de ces numéros spéciaux montre pourtant qu'il existe, au sein même de communauté universitaire et scientifique, au-delà du rejet assez massif de la politique du gouvernement actuel, des différences de sensibilité et un véritable débat, comme l'illustrent bien les éditoriaux très contrastés qui viennent introduire les dossiers de la *Revue d'histoire moderne et contemporaine* et de *Cités*.

L'évaluation : imposture ou véritable enjeu ?

Dans l'éditorial de la revue *Cités*, Yves Charles Zarka voit l'évaluation comme un aspect des « grandes impostures de notre temps, celles qui ont été à l'origine de guerres, de la crise financière et économique gravissime que le monde connaît aujourd'hui », tandis que la *Revue d'histoire moderne et contemporaine* appelle les milieux de la recherche à « prendre à bras le corps [...] et ne pas se laisser dessaisir de cet enjeu majeur qu'est l'évaluation ». Dans le premier cas, il s'agit aussi de dénoncer « la mise en place de dispositifs [...] discrets à leur niveau, très nocifs et même pervers [...]. Un système inquisitorial, qui double et surplombe les procédures existantes d'examen, d'appréciation et de jugement, continue de se mettre en place en dénonçant ceux qui, par hasard, oseraient s'y opposer comme partisans du *statu quo*,

de l'inefficacité et du déclin ». Dans l'autre cas, il s'agit de se mobiliser pour ne pas se laisser « enfermer dans la caricature d'un milieu frileux et arbitraire, qui ne pourrait être ramené que de l'extérieur à une "loi commune" – en vérité totalement hétéronome – de l'"efficacité" et de la "compétitivité" ». En schématisant, on voit bien là transparaître, au-delà des circonstances particulières qui ont présidé à la réalisation de ces deux numéros spéciaux, deux interprétations opposées quant aux relations entre les instances du pouvoir et la société. Il s'agit d'un véritable débat de fond qui traverse les sciences humaines et sociales et qui se traduit par des clivages dans la manière dont les enseignants-chercheurs envisagent leur rapport à la société. Dans un cas, l'évaluation est l'instrument d'un pouvoir qui impose ses représentations d'en haut et qu'il convient de combattre par une critique radicale – un des rôles fondamentaux de l'universitaire est alors de produire des « contre-discours » –, dans l'autre, la société civile – représentée ici par les milieux de la recherche –, est aussi productrice de pouvoir et actrice à part entière dans la formation des normes fondant les relations qui la lient au pouvoir étatique. Les origines de l'évaluation sont ainsi présentées de façon plus ou moins univoque dans les deux revues.

Yves Charles Zarka, dans son article intitulé « L'évaluation : un pouvoir supposé savoir » de la revue *Cités*, voit au cœur du dispositif d'évaluation « une idéologie et un système », l'idéologie étant l'appareil de justification d'un système de contrôle social généralisé et homogène fondé sur une hégémonie du paradigme du marché. En revanche, loin de cette vision monolithique de l'évaluation, l'analyse historique minutieuse que propose Sandrine Garcia dans la *Revue d'histoire moderne et contemporaine* montre que l'introduction de l'évaluation au sein du service public résulte en fait de plusieurs visées bien différentes les unes des autres. Elle s'est initialement inscrite dans une volonté politique de renouveau du service public et de responsabilisation des acteurs politiques. Lors d'un colloque organisé en 1983 par Jacques Delors et Michel Rocard, les tenants de la deuxième gauche pensaient l'évaluation comme « un outil permettant à la société civile de se forger des contre-pouvoirs face à l'expertise légitime des grands corps de l'administration française ». Pour d'autres, elle aurait permis « d'instaurer le pluralisme au cœur de l'expertise de l'action publique et de responsabiliser les acteurs politiques ». Mais à ces deux conceptions s'est opposée une troisième forme, l'évaluation managériale, qui « fonde la validité du jugement évaluatif sur la rigueur scientifique de l'évaluateur et qui place donc la science au-dessus des autres valeurs, en particulier le jugement des acteurs politiques ».

Cependant, au-delà de ces différences d'analyse, les deux revues s'entendent sur le fait que c'est bel et bien la conception managériale qui s'est finalement imposée. Dans la *Revue d'histoire moderne et contemporaine*, Isabelle Bruno montre bien que si l'évaluation se veut de plus en plus quantitative, soumise à des indices de performance, c'est que la recherche est aujourd'hui touchée par des techniques de management issues du monde de l'entreprise dont le *benchmarking* est un des exemples les plus emblématiques. Le *benchmarking* repose en effet sur l'idée que, puisqu'une organisation ne devient compétitive qu'en étant exposée à la concurrence, il incombe à la science managériale d'arranger les conditions d'une compétition interne à l'organisation par la mise au point d'un système d'information confrontant les performances de ses membres. Si cette technique managériale finit par être appliquée au monde de la recherche, c'est dans le cadre d'une progression des idées « néolibérales » depuis les années 1970 qui, dans la mouvance du *New Public Management* – doctrine forgée par les conservateurs britanniques et réappropriée par le gouvernement de Tony Blair – en sont venues à postuler l'analogie entre l'administration d'un État et la gestion d'une entreprise. Et si l'État se doit d'être géré comme une entreprise, la recherche, perçue jusqu'alors comme une des sources de sa puissance, se doit d'entamer une véritable révolution culturelle. C'est cette révolution culturelle que la Commission européenne revendique dans le bulletin d'information de la direction générale chargée de la Recherche : « Le temps où les savoirs acquis dans l'espace scientifique académique constituaient un patrimoine ouvert, mis à la disposition de tous, appartient au passé [...] Le but ultime de la recherche n'est plus simplement de produire des connaissances scientifiques, mais de promouvoir l'exploitation concrète des avancées qu'elle génère. Or cette exploitation, dans une économie de marché, a une dimension intrinsèquement économique»¹. Les nouvelles procédures de l'évaluation en France seraient ainsi la suite logique de la stratégie de Lisbonne qui a fait du *benchmarking* l'instrument principal de la mise en place d'une « économie de la connaissance » européenne.

Usages et mésusages du *benchmarking*

Au style pamphlétaire et lapidaire de la plupart des articles de la revue *Cités*, qui tend parfois à privilégier les développements spéculatifs et les formules choc, on peut préférer l'analyse minutieuse et rigoureuse des origines historiques et de l'environnement institutionnel de cette « fièvre de l'évaluation » que proposent les articles de la *Revue d'histoire moderne et contemporaine*. La revue *Cités* trouve néanmoins un certain nombre de formules justes, comme lorsqu'elle affirme, sous la plume de Philippe Büttgen et Barbara

¹ Commission européenne, « Vers un marché des connaissances », *RDT Info*, n° 34, juillet 2002, p. 16.

Cassin, que l'évaluation « doit servir à classer, c'est-à-dire à justifier des déclassements ou désengagements d'État ». Ainsi, au-delà des discours ronflants sur la compétitivité de la recherche française, il s'agit aussi, tout simplement, de faire des économies. Si la stratégie de Lisbonne se donne la noble ambition de continuer l'intégration européenne en investissant et en transformant le champ de la recherche scientifique pour développer une « économie de la connaissance » européenne, capable d'affronter ses concurrentes américaines et asiatiques, on ne peut que constater dans le même temps l'incapacité ou l'absence de volonté politique d'élever les investissements en R&D. Comme le constate Isabelle Bruno, le *benchmark* des « 3 % », fixé par le Conseil européen de Barcelone, est loin de pouvoir être atteint d'ici 2010. La moyenne des dépenses intérieures de recherche et développement est passée, selon les chiffres d'Eurostat, de 1,92 % en 2000 à 1,91 % en 2006 dans l'Union à 15, et tombe à 1,84 % si les 27 membres sont pris en compte.

Au-delà des doutes que l'on peut avoir sur les intentions réelles des promoteurs des nouvelles formes d'évaluation, la *Revue d'histoire moderne et contemporaine* revient assez longuement sur le caractère inepte et les effets pervers des moyens aujourd'hui utilisés comme instruments de mesure du *benchmarking*. Yves Gingras, dans son article « Du mauvais usage de faux indicateurs », montre que « seule une psychosociologie des dirigeants universitaires et autres fonctionnaires ministériels haut placés pourrait expliquer un tel engouement pour un classement (le classement de Shanghai) qui n'a, en réalité, aucune valeur scientifique ». Un exemple donné par l'auteur parle de lui-même. Le poids démesuré accordé aux prix Nobel a pour résultat de faire « varier la position d'une université de plus de cent rangs dans le palmarès par le seul fait d'attribuer à l'université de Berlin ou à l'université Humboldt le prix Nobel d'Einstein obtenu en 1922 ! » Yves Gingras montre aussi que ce sont souvent les scientifiques eux-mêmes qui succombent aux usages anarchiques de la bibliométrie individuelle. Ainsi, l'indice h, qui est un composite de la production (nombre d'articles écrits) et de la « visibilité » (nombre de citations reçues), peut se justifier en limitant son usage à l'analyse du comportement d'agrégats (université, région, pays), mais est de peu d'utilité au niveau de l'évaluation individuelle étant donné que les « lois de la distribution des publications et des citations ne sont pas des lois normales, mais plutôt log-normales ou encore des lois de puissance de type Pareto ». L'auteur rappelle ainsi qu'Alfred J. Lotka « a montré dès 1926 que la production scientifique suivait une courbe de type Pareto, c'est-à-dire qu'une minorité de chercheurs sont très productifs alors que la majorité des chercheurs

produit peu »². D'autre part, il existe une corrélation très forte entre le nombre d'articles produits par un pays et son niveau de dépense en R&D³. Vouloir réduire la production quantitative d'un chercheur à une question de valeur scientifique individuelle est absurde. Pourtant, cet état d'esprit gagne du terrain et provoque chez les chercheurs des comportements tout à fait contre-productifs, comme le montre Christophe Charle dans la revue *Vingtième Siècle*. L'historien fait remarquer que si l'indice h a remporté un grand succès, c'est parce qu'il « flatte “le narcissisme des scientifiques” (Yves Gingras) en leur donnant l'illusion qu'ils peuvent agir sur leur notoriété et se situer dans une course-poursuite permanente, résumée par un chiffre comme la performance d'un marathonien ». Il ajoute plus loin : « Comme tout historien de la culture le sait, la meilleure manière de se faire remarquer, même si l'on n'est pas remarquable, c'est de faire scandale, de susciter une polémique, de s'attaquer à des sujets controversés, politiquement incorrects, de défendre une thèse paradoxale, etc. ». Il ajoute à cela un autre effet pervers, plus fondamental. Les nouvelles procédures d'évaluation quantitatives pousseraient au conformisme et à la normalisation académique, les chercheurs ne voulant notamment plus publier que dans un nombre limité de revues qui auront été classées.

La question du classement des revues est néanmoins d'importance et peut être envisagée de manière positive, car délimiter une sorte de « périmètre de scientificité » peut être aussi un moyen de protéger le chercheur contre un certain nombre de tentations intrinsèques que l'on a évoquées plus haut et qui menaceraient la qualité et la valeur scientifique de son travail. Malheureusement, comme l'explique Patrick Fridenson dans *Le Mouvement Social*, le classement rendu public fin 2007 par la Fondation européenne de la science, le *European Reference Index for Humanities* (EHIH), a été jugé peu fiable par la British Academy, et trois problèmes ont été identifiés lors d'une session du congrès biennal des historiens allemands. « Le classement ne s'intéresse pas à ce qui devrait être son objet, le travail concret des revues dans la sélection et la mise en forme des articles qui leur sont soumis ainsi que dans la prospection d'articles qu'elle sollicite ; il ne s'appuie pas sur une bibliométrie satisfaisante tenant compte de la diversité des langues et revues européennes, il n'apporte pas l'évaluation dont les revues et les bailleurs de fonds auraient besoin ». Le

² Alfred J. Lotka, « The frequency distribution of scientific productivity », *Journal of the Washington Academy of Science*, 36, 1926, p. 317-323.

³ E. Archambault, D. Campbell, Y. Gingras, V. Larivière, « WOS vs Scopus: on the reliability of scientometrics », *Books of Abstracts of the 10th International Conference on Science and Technology Indicators*, 2008.

classement de l'AERES a lui été réalisé sur la base d'une méthodologie assez confuse, de « bric et de broc », pour reprendre l'expression lapidaire de Patrick Fridenson.

Les nouvelles procédures d'évaluation sont mises en place dans la précipitation. Là aussi, la recherche semble soumise à une communication politique du résultat à court terme qui explique sans doute en partie le rejet de la politique du gouvernement par les milieux universitaires et de la recherche. Mais, au-delà de l'émotion commune face à ce qui est ressenti comme une attaque contre la manière dont les enseignants-chercheurs conçoivent l'apport singulier de la science à la société, le contraste entre les lignes éditoriales de *Cités* et de la *Revue d'histoire moderne et contemporaine* montre aussi qu'il existe des divergences assez profondes sur la question de l'évaluation. Et c'est dans l'animation de ce débat-là que les revues de sciences humaines et sociales, conjointement à leur vocation scientifique, nous sont aujourd'hui précieuses.

Publié dans www.laviedesidees.fr, le 28 mai 2009

© laviedesidees.fr