

Que valent les évaluations ?

par Coline Soler

Les universités et la recherche font l'objet de réformes et d'évaluations depuis vingt ans. Clémentine Gozlan étudie la fabrique de ces dispositifs dans le nouveau gouvernement de la science auquel nombre de chercheurs eux-mêmes donnent la main.

À propos de : Clémentine Gozlan, *Les valeurs de la science. Enquête sur les réformes de l'évaluation de la recherche en France*, ENS Éditions, 266 p., 28 €.

De nombreuses réformes visent à transformer le système d'enseignement supérieur et de recherche depuis les années 2000. Celles-ci ont donné lieu à autant de travaux concernant les transformations des modes de gouvernement des professions, les recompositions au sein des groupes professionnels, les effets de l'évaluation sur la production des savoirs et les manières dont les universitaires s'approprient les classements et les indicateurs. Face à cela, l'originalité et l'intérêt de l'ouvrage de Clémentine Gozlan sont de discuter le rôle des chercheurs et chercheuses dans la conception des instruments managériaux et politiques et d'étudier la fabrique des dispositifs plutôt que de se concentrer sur les effets des réformes.

L'évaluation est une activité familière pour les chercheurs et chercheuses, inscrite dans les pratiques scientifiques : il s'agit aussi bien de l'évaluation des publications par les pairs que de l'évaluation des laboratoires et centres de recherche par des instances spécifiques. Pourtant, la création de l'Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES) en 2007 a suscité force débats et réticences dans la communauté scientifique, que l'autrice explique par les changements majeurs introduits dans le gouvernement de la science.

Partant de ce constat, l'auteur propose d'étudier cette institution comme un observatoire des controverses et des luttes de pouvoir dans la communauté scientifique, un lieu où se construisent les normes professionnelles et les pratiques de jugement, pendant ses cinq années d'existence (2007-2012). Pour ce faire, l'enquête se focalise sur une des trois sections de l'AERES (celle s'occupant de l'évaluation des laboratoires ou unités de recherche) et sur les sciences humaines et sociales (SHS), terrains moins traités par les travaux d'histoire et de sociologie des sciences (Lamont, 2009). Dans ces disciplines, l'introduction de nouveaux modes d'évaluation a été la plus controversée (Aust, Gozlan, 2018), ce que l'auteur explique par le fait que certains instruments d'évaluation vont à l'encontre des conceptions de la qualité en recherche (par exemple, le financement sur projets ou les classements). L'enquête consiste en des entretiens approfondis, des observations directes de commission, complétée par des analyses (qualitatives et quantitatives) de nombreuses sources écrites.

Les enjeux soulevés par cet ouvrage sont multiples, nous en détaillerons ici trois : la conception des outils d'évaluation, leur contestation et l'appropriation qu'en font les évaluateurs.

Construction de l'agence et conception des outils d'évaluation

L'enquête permet d'aller au-delà de l'idée selon laquelle l'AERES serait le résultat d'une réforme imposée « par le haut » aux universitaires, bien que la création de l'agence soit une initiative politique. Même si de nouveaux acteurs extérieurs à la profession scientifique (gestionnaires, hauts fonctionnaires, consultants, etc.) sont impliqués, l'auteur montre que les enseignants-chercheurs ont été centraux dans ce processus, ce qui lui permet d'affirmer que l'agence n'est ni « une courroie de transmission de l'État qui à travers elle gouverne la communauté scientifique. » (p. 42), ni entièrement assujettie au contrôle politique.

L'apport de cette démonstration est de dépasser l'opposition interne/externe à la profession ou encore l'alternative scientifique/manager et de décrire avec précision les relations entre différents acteurs internes à l'AERES (administratifs, scientifiques,

experts qualité¹) et leur capacité d'influence variable sur la définition des procédures d'évaluation.

Une représentation des trajectoires de quarante délégués scientifiques en SHS et de différents délégués coordinateurs en poste entre 2007 et 2011, grâce à une analyse des correspondances multiples (ACM)², permet de mettre en évidence leurs ressemblances et leurs dissemblances. L'autrice identifie trois profils : le « scientifique » (faible participation aux instances gouvernantes de la profession), le « responsable universitaire » (responsabilités à l'échelle des universités, centre de recherche ou discipline), le « responsable politico-administratif » (responsabilité à l'échelle nationale ou politique). Ceux qui participent à la construction des procédures d'évaluation sont des collègues plus que des dirigeants : ce ne sont ni des experts externes au monde scientifique ni des ex-pairs qui se sont éloignés durablement de la profession scientifique et de la base professionnelle. L'autrice les qualifie « d'élites intermédiaires », ce qui illustre leur position à mi-chemin entre l'élite dirigeante et les professionnels de la base, car ils continuent d'être chercheurs ou enseignants-chercheurs, en pratique et dans leurs représentations de soi.

L'autrice s'interroge ensuite sur les raisons qui poussent une partie des universitaires à vouloir participer à la conception des outils de jugement de l'AERES (l'ensemble des standards, normes, chartes et codes de bonne conduite pour l'évaluation, de la préparation du comité de visite jusqu'à la rédaction du rapport d'évaluation). Après avoir réinscrit la bureaucratisation des politiques scientifiques dans un temps plus long, deux enjeux semblent expliquer ce mouvement de rationalisation. D'une part, il répond aux discours critiques sur les instances d'évaluations existantes, suspectées de proximité voire de conflits d'intérêts. D'autre part, la volonté de mettre en place des standards valides quelles que soient les spécificités disciplinaires, pour la production de rapports d'évaluation homogènes et comparables renvoie à une conception procédurale de la justice qui implique que « l'équité du jugement [soit] assurée par le respect scrupuleux de procédures transparentes et identiques pour tous » (p. 82).

¹ Une « cellule qualité » est créée début 2008 à l'AERES pour accompagner la mise en place des procédures d'évaluation afin de les rendre conformes aux standards européens, dans l'objectif de voir l'agence reconnue comme membre de l'European Association for Quality Assurance in Higher Education.

² Il s'agit d'une méthode statistique de traitement de données permettant de synthétiser des tableaux qui renseignent des caractéristiques qualitatives pour une population d'individus et de représenter les informations les plus structurantes sous forme de graphiques.

Toutefois, malgré la volonté des présidents de l'agence de mettre en place ces standards, le travail concret des membres de l'AERES consiste plutôt en une forme de « bricolage », un ensemble de tâtonnements et de réflexions internes. Cette bureaucratisation s'accompagne de la fabrication d'une conception de la « bonne manière d'évaluer » (p. 80), c'est-à-dire de la capacité des évaluateurs à respecter les règles et normes préétablies, qui n'est pas toujours bien accueillie.

Réinvention et contestation des outils d'évaluation

Avec deux études de cas, l'autrice met en évidence une capacité inégale de contestation des instruments qui dépend de plusieurs facteurs : caractéristiques des contestataires, capacité à construire une position protestataire commune, à solliciter des relais légitimes, caractéristiques des instruments eux-mêmes, proximité avec les représentations et pratiques professionnelles.

Le premier cas est la mobilisation contre le classement des revues par les représentants à l'AERES de la recherche en littérature (présidents d'associations savantes, de sections du Comité National des Universités, du Comité national du CNRS, directeurs des revues du champ disciplinaire concerné, certains délégués scientifiques à l'AERES). Plusieurs ressorts de cette critique sont développés : en relativisant l'objectivité et la légitimité des classements, ces représentants cherchent à re-politiser l'instrument et à mettre en évidence ses multiples biais ; en donnant à voir les particularités de la discipline, ils s'opposent aux classements et revendiquent leur autonomie professionnelle. Ce travail discursif ne suffit pas au succès de l'opposition, qui repose aussi sur « les rapports de pouvoir entre les institutions concurrentes de gouvernement de la profession » (p. 129). Les représentants de la discipline revendiquent une expertise propre et parviennent à mobiliser des relais influents pour contester la mise en place des classements, aidés par des instances représentatives (les associations savantes comme la Société française de littérature générale et comparée, par exemple) qui sont devenues « incontournables dans la fabrique des politiques d'enseignement et de recherche » (p. 133), ne serait-ce que parce que certaines d'entre elles suggèrent les programmes de l'agrégation.

La deuxième étude de cas se penche sur la mise en place en 2011 d'une consultation interne de plusieurs mois, à laquelle l'autrice a participé et qui débouche sur la rédaction d'un référentiel d'évaluation définissant « les conditions de

scientificité auxquelles les écrits des chercheurs doivent souscrire pour être considérés comme scientifiques » (p. 143). L'enquête rend compte du processus de construction des normes d'évaluation, de la définition de la science en train de se faire, de la valeur scientifique et finalement de « ce que "faire de la recherche" veut dire » (p. 147). Un premier travail de définition des frontières de la production scientifique aboutit à la distinction entre science et non-science (vulgarisation, création, engagement militant). La posture adoptée consiste à ne pas juger du fond, mais à construire un instrument permettant une évaluation formelle et procédurale de la scientificité, actant la reconnaissance de l'hétérogénéité des règles de valorisation de la production académique dans les disciplines de SHS. Cela permet d'affirmer la position de l'AERES en dehors des controverses scientifiques propres à chaque discipline, pour ne pas favoriser une pratique ou une conception de la science au détriment d'autres, plus minoritaires ou marginales. L'autrice note toutefois que même si le référentiel ne juge pas de la qualité des résultats de recherche, la formalisation du protocole agit tout de même sur les pratiques de publication et « sur les lois de la reconnaissance scientifique » (p. 176).

Appropriation des outils d'évaluation

Enfin, le processus d'évaluation et ses effets performatifs sont analysés en comparant deux disciplines des SHS, les études littéraires et la géographie, à l'aide de plusieurs méthodes (étude quantitative des rapports AERES, entretiens, sources écrites, observation d'un comité de visite, c'est-à-dire du groupe de pairs qui se rend dans les unités et équipes de recherches pour rencontrer les personnels et évaluer leur travail). Cette comparaison est féconde, car elle permet de montrer l'appropriation différente des instruments : si les littéraires discréditent les outils de l'AERES, les géographes en font un usage beaucoup plus systématique. L'autrice explique cela notamment par la position des disciplines dans le monde académique : la géographie est représentée comme une discipline « dominée » (Bourdieu, 1984, p. 145), alors que la littérature est au pôle opposé de l'échelle du prestige en SHS (Bourdieu, 1984). La discipline est ainsi dotée d'une plus forte reconnaissance en France, comparativement aux États-Unis (Duell 2000), l'enquête ne montrant pas de signes d'une crise de légitimité comme celle décrite par les enseignants-chercheurs en géographie. Un autre facteur explicatif tient à la façon dont les réformes sont perçues comme allant dans le sens de l'intérêt de la discipline ou non : les injonctions au travail collectif et aux

publications internationales allant dans le sens de l'organisation du travail en géographie, mais pèsent sur les modes de travail en littérature.

Conclusion

L'autrice conclut que la mise en place de l'AERES a entraîné une « redistribution au sein de la communauté académique du pouvoir d'énoncer les normes qui comptent » (p. 224). Les outils de l'AERES réforment « à distance » les pratiques scientifiques puisqu'ils ciblent davantage les procédures que la qualité scientifique.

Cet ouvrage propose des éléments de compréhension des débats qui ont accompagné la création de l'AERES et n'intéressera pas seulement les chercheurs travaillant sur ces questions, car il ouvre plus largement de nombreuses réflexions sur des enjeux importants pour comprendre et penser le système de recherche, d'enseignement supérieur et le travail scientifique. Bien que l'enquête porte sur la période 2007-2012, les questions qu'elle soulève sont d'autant plus d'actualité que le nouveau référentiel du Haut Conseil de l'Évaluation de la Recherche et de l'Enseignement Supérieur (HCERES, qui a pris la suite de l'AERES) présenté en novembre a suscité lui aussi une vive réaction de l'Assemblée des directions de laboratoires qui dénonce une bureaucratisation extrême au détriment de la collégialité et du projet scientifique des unités évaluées.

Publié dans laviedesidees.fr, le 19 janvier 2022