

INTRODUCTION

La création de l'AERES a «secoué le cocotier», en faisant passer tout le monde dans la logique de l'évaluation et du «rendre-compte». Le principe d'une évaluation externe est entré dans les mœurs, même dans les disciplines où ce n'était pas forcément très présent.¹

Depuis les années 2000 en France, le monde de l'enseignement supérieur et de la recherche a connu une série de réformes sans précédent. De nouvelles règles du jeu modifient les structures de financement de la recherche, l'organisation des universités, le travail des chercheurs, et l'évaluation scientifique². À l'automne 2007, dans un contexte de fortes mobilisations contre l'une des lois transformant le paysage universitaire, la loi relative aux libertés et responsabilités des universités (LRU), l'Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES) entre en action. « Bras armé du ministère », « œil de Moscou », ou encore « suppôt de Satan » – pour reprendre les termes employés par certains universitaires –, l'AERES a fait l'objet de dénonciations virulentes, décrivant souvent un tournant néo-managérial qui réduirait l'autonomie professionnelle. Les ardentes critiques qui ont marqué la création de cette agence peuvent à première vue paraître étranges : depuis longtemps, l'évaluation de la recherche fait partie intégrante des activités des universitaires et des chercheurs, qui, à différents moments de leur

1 Frédéric Dardel, président de l'université Paris 5, 2014. En ligne : [<http://www.letudiant.fr/educpros/enquetes/le-rapport-des-universitaires-sur-l-agence-d-evaluation-de-la-recherche-et-de-l-enseignement-superieur-aeres.html>].

2 Pour une analyse de ces transformations, voir notamment Musselin (2017).

carrière, sont évalués et expertisent les travaux de leurs collègues. Le travail d'évaluation et de validation collégiale des avancées scientifiques est un pilier de la régulation de la profession académique.

Dès le XVIII^e siècle, les publications de recherche commencent en effet à être soumises à l'évaluation « par les pairs » ou *peer review* (Gingras, 2014). Ce terme désigne l'activité collective de jugement critique, par des chercheurs, des travaux de leurs collègues. Ainsi, les résultats d'une recherche ne peuvent être publiés dans une revue scientifique que s'ils ont été examinés et entérinés par des spécialistes appartenant à la même profession et, généralement, à la même discipline. Les chercheurs sont donc évalués individuellement, dès qu'ils soumettent leurs travaux à des revues scientifiques ou à des comités éditoriaux. Mais les publications ne sont pas les seuls objets d'évaluation : en France, au fur et à mesure de la structuration de la recherche au long du XX^e siècle, se mettent en place des instances visant à évaluer les organisations où travaillent les scientifiques – les laboratoires ou centres de recherche. Ainsi, la naissance du Centre national de la recherche scientifique (CNRS) en 1939 est presque immédiatement suivie par la mise en place du Comité national de la recherche scientifique (CoNRS) du CNRS, créé à la fin de la guerre pour évaluer les individus et les équipes sous tutelle de cet organisme de recherche³.

L'évaluation de la recherche n'est donc pas une nouveauté au moment de la création de cette agence. Elle est même une activité routinière, familière aux chercheurs et fortement inscrite dans leurs pratiques et leur quotidien de travail. Pour comprendre pourquoi la mise en place de cette nouvelle instance a provoqué autant de controverses et de débats dans la communauté scientifique, il faut souligner les changements majeurs qu'elle véhicule dans le gouvernement de la science.

PÉRIL EN LA DEMEURE : UNE RÉFORME MASSIVE DE L'ÉVALUATION SCIENTIFIQUE ?

Le premier changement introduit par l'AERES réside dans les finalités symboliques et financières de l'évaluation. La création de cette agence s'inscrit dans une séquence de réformes inédites de l'enseignement supérieur et de la recherche (Ravinet, 2012). Compétition pour les crédits scientifiques et évaluation de la performance des universitaires sont les

3 Pour un aperçu des structures d'évaluation scientifique en France, voir l'annexe 1.

pièces maîtresses des réformes qui voient le jour dans le Pacte pour la recherche de 2006. Accompagnant les discours sur le « retard français » (Bouchard, 2008) dans la course à l'excellence, le nouveau système évaluatif est supposé rendre plus efficace la détection du mérite académique, en indexant notamment une partie des crédits des organisations universitaires sur les résultats de l'évaluation, comme dans d'autres pays, tels que la Grande-Bretagne⁴. Dès lors, l'évaluation devient un instrument de pilotage à la performance et elle est perçue comme un vecteur potentiel d'accroissement des inégalités au sein du monde scientifique (Musselin, 2017), selon l'« effet saint Matthieu » (Merton, 1968)⁵.

La création de l'AERES s'accompagne ensuite de la promotion d'outils de mesure et de classement. Les membres de cette agence attribuent des notes aux centres de recherche et calculent des indicateurs, tels que les taux de « publiants », c'est-à-dire de chercheurs et enseignants-chercheurs jugés actifs en recherche, en fonction du nombre d'articles qu'ils publient. Les réactions ne tardent pas à fuser dans la communauté académique : cette entreprise de chiffrage est suspectée de profondément remettre en cause les pratiques qualitatives de jugement caractéristiques de l'évaluation par les pairs, fondée sur la lecture des travaux de recherche, au profit d'une définition comptable de la qualité scientifique. Cependant, l'utilisation de notes et d'instruments de classement n'est pas introduite par l'AERES⁶. C'est surtout la publicité nouvelle donnée à ces outils d'évaluation qui cristallise les critiques. En effet, depuis la création de l'AERES, l'intégralité des rapports d'évaluation et des notes attribuées aux laboratoires est rendue publique, alors que ce n'était pas le cas des évaluations conduites par d'autres instances auparavant. Cette visibilité peut transformer le sens et l'usage des évaluations : outils de stigmatisation potentielle – si les notes sont mauvaises –, elles deviennent des outils

4 La rémunération des universitaires à la performance y a été mise en place plus tôt et de façon plus significative qu'en France : créé en 1986 sous le gouvernement de Margaret Thatcher, le *Research Assessment Exercise* (RAE) détermine le volume des fonds publics attribués aux départements de recherche en fonction des résultats de leur évaluation.

5 Ce terme désigne le fait que ceux qui ont déjà une reconnaissance scientifique concentrent les ressources. Pour le forger, Robert K. Merton s'est inspiré d'une phrase de l'Évangile selon Matthieu : « Car on donnera à celui qui a, et il sera dans l'abondance, mais à celui qui n'a pas on ôtera même ce qu'il a ».

6 Certaines sections du Comité national du CNRS pratiquaient par exemple déjà la notation des laboratoires, tandis que le taux de « publiants » a été inventé au sein du ministère de la Recherche dans les années 1990. Nous reviendrons dans le premier chapitre sur l'histoire de cet indicateur contesté.

de *benchmarking* (Bruno et Didier, 2013), c'est-à-dire des technologies de gouvernement fondées sur la comparaison de tous contre tous, incitant les chercheurs à répondre aux injonctions implicites contenues dans les critères d'excellence pour se positionner dans cette course à la performance. L'évaluation devient ainsi un instrument central d'orientation et de gouvernement des conduites des universitaires.

Enfin, la centralité et l'étendue des missions de cette agence – évaluer l'ensemble du dispositif de recherche et d'enseignement supérieur, au moyen de procédures et de critères unifiés – modifient les conditions de l'expertise collégiale et les critères de qualité scientifique. En effet, l'évaluation de la recherche était jusqu'alors conduite par des commissions disciplinaires, ou, dans certains cas, pluridisciplinaires, qui décidaient des normes légitimes pour juger la recherche en fonction des particularités des disciplines représentées et de leurs modes de production des savoirs. Or, avec la création de l'AERES, l'évaluation est désormais confiée à une institution centrale dont le mandat, inscrit dans la loi, est de proposer des procédures identiques pour toutes les disciplines et les types de laboratoires de recherche. Ce projet de standardisation et d'uniformisation n'est pas sans poser problème : les critères d'appréciation de la qualité scientifique sont fortement hétérogènes selon les domaines (Lamont, 2009), et la consécration scientifique est bien souvent un objet de luttes entre des écoles de pensée, entre des disciplines, entre des collègues (Barber, 1961; Bourdieu, 1975, 1984, 2001a; Latour, 1984). Dans la mesure où uniformiser les outils d'évaluation, c'est imposer, *a priori*, une définition dominante des pratiques de recherche et des outils pertinents de leur évaluation, les débats sur la valeur académique trouvent une actualité renouvelée avec la création de l'AERES.

Cet ouvrage étudie donc une activité centrale et récurrente au sein de la communauté scientifique – l'évaluation – à un moment particulier où ses règles sont bouleversées, où ses modalités sont repensées, où de nouveaux acteurs sont susceptibles de prendre le pouvoir sur les critères employés pour juger les activités de recherche. L'AERES devient ainsi un observatoire pertinent des transformations des manières de juger la valeur du travail scientifique, des controverses et des luttes qu'elles suscitent.

L'ouvrage propose moins une réflexion institutionnelle sur cette agence qu'il ne la saisit comme un microcosme où sont débattues et réinventées les normes professionnelles – qu'est-ce qu'un « bon » universitaire ? Quelles sont les pratiques de recherche légitimes ? Comment évaluer le travail des chercheurs ? Au moyen d'une enquête de terrain auprès des membres de cette agence et d'autres acteurs du monde académique, il s'agira de retracer les controverses à l'issue desquelles se construisent certaines définitions du travail scientifique et des pratiques d'attribution de la valeur.

Pour saisir ces débats, je me centre sur les aspects les plus conflictuels des transformations de l'évaluation de la recherche. L'ouvrage porte sur le « moment AERES », à savoir ses cinq années d'existence (2007-2012), avant que, face aux multiples critiques émanant de la communauté scientifique, elle ne soit remplacée par une autre structure⁷. Cette période de création de normes est marquée par des affrontements particulièrement virulents au sein de la profession scientifique, entre des représentants de différentes disciplines et de différentes institutions, que la focale sur ce moment réformateur permet d'explorer.

Dans cette perspective, je m'intéresse plus spécifiquement à l'une des trois sections de l'AERES, celle chargée de l'évaluation des « laboratoires » ou « unités de recherche »⁸. L'évaluation de la recherche à l'AERES ne porte en effet pas sur des individus, mais sur l'activité de structures de recherche placées sous la responsabilité d'établissements d'enseignement supérieur et/ou d'organismes comme le CNRS. Placer la focale sur cette section présente deux intérêts. D'une part, cette section illustre par excellence le mouvement de centralisation de l'évaluation lancé par l'AERES, puisqu'elle reprend des missions auparavant confiées à des instances bien distinctes, notamment le CoNRS et la Mission scientifique, technique et pédagogique (MSTP) du ministère de l'Enseignement supérieur et de la Recherche (voir tableau 1). On peut alors retracer en quoi l'AERES s'inspire des pratiques et instruments d'évaluation qui avaient été déjà

7 Il s'agit du Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur (HCERES), mis en place en 2013. Il conserve la mission d'évaluation tout en en faisant évoluer les procédures.

8 Les deux autres sections de l'AERES sont respectivement chargées de l'évaluation des établissements et de l'évaluation des formations supérieures. Voir l'annexe 4, qui présente et décrit l'organigramme de l'AERES.

développés au sein de ces instances, mais aussi voir comment elle s'en démarque. D'autre part, en étant chargée de l'évaluation de l'ensemble des unités de recherche, toutes disciplines confondues, cette section est un observatoire idéal pour comprendre comment se construisent des procédures et des normes d'évaluation transversales.

Structure de recherche	Définition	Évaluation avant la création de l'AERES
Unités de recherche labellisées par le CNRS	Structures de recherche reconnues et financées par le CNRS. Elles sont composées de chercheurs CNRS, auxquels peuvent s'ajouter des chercheurs d'autres organismes et des enseignants-chercheurs rattachés à un établissement d'enseignement supérieur.	Avant la création de l'AERES, elles sont évaluées par le Comité national de la recherche scientifique, une instance composée de membres de la communauté académique, créée en 1945 et placée sous la responsabilité du CNRS. Pour ne pas être confondue avec le CNRS, cette instance est souvent désignée simplement comme le Comité national ou, à l'écrit, comme le CoNRS. Même si une partie de ses missions est reprise par l'AERES, cette structure reste en place après sa création.
Les « équipes d'accueil » reconnues par le ministère de l'Enseignement supérieur et de la Recherche	Structures de recherche reconnues et soutenues financièrement par le ministère de l'Enseignement supérieur et de la Recherche, mais ne bénéficiant pas de la labellisation par un organisme de recherche. Composées uniquement d'enseignants-chercheurs, elles sont qualifiées d'« équipes d'accueil » parce qu'elles peuvent accueillir des étudiants en thèse.	Avant la création de l'AERES, elles sont évaluées par la Mission scientifique et technique (MST), une structure rattachée au ministère de l'Enseignement supérieur et de la Recherche. Cette instance, qui existait, depuis les années 1980 et a pris d'autres dénominations au cours du temps, telles que Mission scientifique, technique et pédagogique (MSTP) et Mission scientifique universitaire (MSU). Elle disparaît avec la création de l'AERES.

Tableau 1. L'évaluation des unités de recherche lors de la création de l'AERES

Pour mener à bien ce questionnaire, j'ai choisi de me centrer sur les disciplines dans lesquelles l'introduction de nouveaux modes d'évaluation a été la plus controversée, à savoir les sciences humaines et sociales (SHS). Ces disciplines ont souvent été délaissées par les travaux d'histoire et de sociologie des sciences. À de rares exceptions près (Lamont, 2009), ces travaux sont plus volontiers centrés sur les sciences expérimentales (Martin, 2006) ; l'invisibilisation des sciences humaines dans les *science studies* pouvant être considérée comme une trace de leur déficit de légitimité scientifique et sociale. Ces disciplines sont également réputées indis-

ciplinées, ou du moins, rétives à toute forme de standardisation (Becher, 1994 ; Becher et Trowler, 2001). Certes, les résistances concernant l'usage d'outils d'évaluation standardisés débordent les SHS⁹, mais c'est dans ces disciplines que les protestations et les mobilisations ont été les plus vives (Aust et Gozlan, 2018). En effet, certains instruments d'évaluation promus au sein de l'AERES heurtent frontalement les conceptions de la qualité en recherche en SHS, et il s'agit d'étudier comment les universitaires de ces disciplines s'accommodent de ces instruments ou au contraire les délégitiment, voire comment ils s'investissent dans la production d'outils de jugement alternatifs. Les débats que posent les transformations de l'évaluation en SHS font figure de miroir grossissant des luttes de pouvoir, parfois plus discrètes, qui traversent le champ académique et qui ont pour enjeu tant les définitions de la recherche scientifique que les formes légitimes d'expertise dans la profession.

Le titre de cet ouvrage, *Les valeurs de la science*, renvoie d'une part à l'activité de construction de conventions relatives à ce qui compte en termes de production scientifique, c'est-à-dire à la valorisation ou à la dévalorisation de certaines activités de recherche, à travers les critères de jugement qui sont produits à l'AERES. Dans cette perspective, il prête une attention particulière au travail normatif mené par les membres de l'agence, c'est-à-dire la production de normes, de critères et de procédures destinés à encadrer les évaluations réalisées par l'AERES. D'autre part, il met en avant la pluralité des modes d'attribution de la reconnaissance scientifique et des échelles de valeurs. En effet, les chercheurs ne s'entendent pas tous sur les bonnes façons de mesurer la qualité du travail académique, ni même sur la possibilité d'une telle mesure. En fonction de leurs trajectoires, de leurs disciplines ou de leur positionnement scientifique, ils peuvent embrasser des axiologies, des systèmes de valeurs profondément différents. Ce livre entend apporter, à son échelle, quelques éléments pour mieux comprendre les débats sur la science, sa place et sa valeur dans notre société.

9 Les SHS ne sont pas les seules disciplines à faire porter leur voix contre des instruments qu'elles estiment inadéquats. En témoigne par exemple la position critique de Richard R. Ernst, prix Nobel de chimie en 1991, envers la prégnance des classements (Ernst, 2010). À l'AERES, les débats sur les façons de juger la qualité mettent également aux prises les représentants des sciences fondamentales et ceux des sciences appliquées, qui militent, non sans difficulté, pour que la notion de « production scientifique » soit ouverte à d'autres types de travaux que ceux publiés dans des revues à comité de lecture, et pour que l'« impact social » et économique de leurs recherches soit reconnu dans l'évaluation par l'agence. Pour une analyse des controverses sur la place sociale de la science à l'AERES, voire notamment Gozlan (2015).

En étudiant les transformations de l'évaluation de la science, cet ouvrage contribue aux débats foisonnants sur la mise en place de réformes managériales et leurs effets sur les mondes professionnels. Depuis les années 2000 en France se sont succédé des réformes visant à accroître l'efficacité des agents, souvent au moyen d'instruments de mesure quantifiée de leur performance au travail. Le monde scientifique n'échappe pas à ces tendances : pour plusieurs auteurs, les réformes du secteur académique sont directement inspirées des préceptes du *New Public Management* (NPM) qui se diffuse dans le monde anglo-saxon depuis les années 1980 (Vinokur, 2008). « Le recours systématique aux mécanismes de marché, [et] la mise en place d'une gestion par les résultats fondée sur la réalisation d'objectifs, la mesure et l'évaluation des performances et de nouvelles formes de contrôle » (Bezès et Demazière, 2011, p. 296) sont deux caractéristiques centrales du NPM (Bezès, 2009, 2012 ; Pollitt, 1990).

À première vue, l'introduction de ces mécanismes de contrôle heurte très directement les modes de fonctionnement des professions établies¹⁰. Dans la littérature anglophone, le terme de « professionnel » désigne des travailleurs dotés d'une forte expertise, qui revendiquent une grande autonomie à la fois dans la conduite quotidienne et dans la régulation collective de leurs activités (pour le cas des chercheurs, voir Barrier, 2011). Cette autonomie se manifeste dans le fait que les membres de la profession forment et cooptent leurs futurs collègues, et se voient reconnaître une compétence et une expertise spécifiques dans la conduite de leurs activités. En effet, une part importante du travail de défense de leur statut dépend de leur capacité à revendiquer la maîtrise d'un savoir qu'ils sont seuls à détenir (Paradeise, 1985). La revendication du monopole de la compétence à juger constitue un trait central de la rhétorique professionnelle, selon laquelle « [la] profession est seule dépositaire de compétences techniques [...], seuls les détenteurs de cette compétence sont capables de juger du caractère vrai ou faux, bon ou mauvais, de celle-ci » (Damien et Tripiet, 1994, p. 217). Dans ce modèle typique d'autorégulation, le rôle de

¹⁰ Pour Michael Power (1997) par exemple, « les programmes relevant du *New Public Management* ont pour objectif manifeste un certain degré de colonisation. [...] Leur intention [...] est de défier le pouvoir et la liberté d'action de groupes relativement autonomes en les rendant publiquement responsables de leurs performances » (p. 184-185).

l'État se borne à garantir et à maintenir l'autonomie des professionnels (Freidson, 1984a, 1984b, 2001; Champy, 2009).

Aux côtés des médecins et des avocats, la profession scientifique constitue un cas emblématique du modèle professionnel. Certains auteurs voient même dans l'autonomie scientifique une condition nécessaire à la poursuite de l'activité de recherche. Michel Polanyi (1962), défendant la « République de la science », considère par exemple que « toute autorité qui s'engagerait à diriger et à orienter le travail des chercheurs freinerait le progrès de la science » (p. 56)¹¹. L'organisation collégiale de la science apparaît comme seule garante de l'avancée des travaux académiques : la nature du travail scientifique requiert la pratique d'un savoir informé et spécialisé, qu'aucun acteur externe à cette communauté ne saurait mettre en œuvre, ni *a fortiori* juger. La création de l'AERES, nouvelle autorité centrale d'évaluation, questionne ce modèle d'autorégulation et d'autonomie professionnelle. Autorité administrative indépendante (AAI)¹², l'AERES rend néanmoins des comptes – comme la plupart des agences – au ministère de la Recherche et de l'Enseignement supérieur, semblant ainsi accroître le contrôle étatique sur les universitaires et les chercheurs. Par ailleurs, l'affichage de notes et d'indicateurs « met en grille » le jugement (Cassin, 2014) et concurrence l'expertise professionnelle qualitative. La création de cette agence semble bouleverser les conditions de l'autonomie des universitaires, et il s'agit d'étudier dans quelle mesure.

Cet ouvrage s'inscrit ainsi à la suite des travaux sur les transformations des modes de gouvernement des professions. En effet, les réformes des mondes professionnels – qu'il s'agisse de la médecine (Pierru, 2007), de la justice (Mouhanna, 2015), de la police (de Maillard, 2009) – ont ouvert un champ de recherche particulièrement fécond donnant lieu à de vifs débats, auxquels cette enquête entend contribuer. Un premier ensemble de travaux considère que les réformes néo-managériales assujettissent les professionnels, mettent en péril leur autonomie (Belorgey, 2010) et conduisent à leur « déprofessionnalisation » (de La Broise, 2013; Mas *et al.*,

11 Il y aurait beaucoup à dire sur les usages « stratégiques » de l'autonomie professionnelle, en fonction du contexte social et politique. C'est en effet en 1942 que Robert K. Merton publie son article « The Normative Structure of Science » (réédité ultérieurement, Merton, 1973, p. 267-278) : souligner, en période de guerre, que le progrès scientifique obéit à des normes professionnelles spécifiques, c'est aussi s'opposer à une conduite de l'activité scientifique marquée par l'interventionnisme étatique et la mise sous contrôle idéologique.

12 Article L. 114-3-1 de la loi de programme du 18 avril 2006 pour la recherche. Son statut implique qu'elle n'est pas soumise à l'autorité hiérarchique de représentants de l'État.

2011). D'autres études soulignent surtout les rapports de concurrence entre les professionnels et de nouveaux acteurs jusque-là discrets dans le gouvernement du secteur considéré. L'une des transformations contemporaines majeures réside dans le passage d'une définition des pratiques et normes professionnelles « de l'intérieur », par les professionnels eux-mêmes, à la prégnance accrue d'injonctions portées « de l'extérieur » des professions sur la conduite de leurs activités, bien que l'articulation entre ces formes de contrôle soit variable (Boussard, Demazière et Milburn, 2010). Si l'immixtion de nouveaux *managers* dans la définition du fait professionnel ne s'accompagne pas toujours d'un recul de l'autonomie (Robelet, 2001; Hassenteufel, 1999; Farrell et Morris, 2003), elle révèle que le pouvoir des professionnels est fragilisé, en recomposition.

Mettant à l'épreuve l'opposition entre autonomie « traditionnelle » des professions et hétéronomie, un troisième ensemble d'études pose davantage le problème en termes de mutations du professionnalisme¹³ (Evetts, 2003; Bezès et Demazière, 2011). Plusieurs travaux ont étudié les recompositions qui opèrent au sein même des groupes professionnels. Les réformes sont une occasion de redistribution du pouvoir et de création de nouveaux pouvoirs (Bezès, 2007). Elles peuvent ouvrir à certaines franges professionnelles l'opportunité de conquérir une autorité sur d'autres (Freidson, 1994; Robelet, 2001), tandis que la rationalisation de l'activité peut faire l'objet de luttes horizontales au sein même de la profession (Castel et Merle, 2002; Bergeron et Castel, 2014).

Ce débat sur les liens entre professions et nouveau management public (Bezès et Demazière, 2011) est particulièrement pertinent pour étudier les transformations du monde scientifique. Diverses enquêtes ont ainsi exploré les conséquences plurielles des réformes sur le monde scientifique. Le cas français n'est pas isolé : plusieurs travaux montrent ainsi comment, Outre-Manche, le RAE redistribue le pouvoir entre les départements scientifiques (Camerati, 2014), transforme la gouvernance des universités britanniques (McNay, 1997), concentre les crédits sur les universités les plus prestigieuses et renforce ainsi les inégalités scientifiques existantes (McNay, 2003). D'autres relativisent la portée des réformes des politiques scientifiques : Mary Henkel (2002) montre que malgré les

13 Cette voie a été ouverte par Eliot Freidson (1994) qui soutenait « que les éléments essentiels du professionnalisme ne sont pas en train de disparaître, mais qu'ils prennent plutôt une nouvelle forme. Le professionnalisme est en train de renaître sous une forme hiérarchique dans laquelle les praticiens sont soumis au contrôle des élites professionnelles » (p. 9).

modifications de leur environnement de travail et le contrôle accru qui pèse sur leurs pratiques, l'identité et les valeurs académiques des universitaires restent relativement stables. D'autres auteurs s'intéressent aux effets des instruments d'évaluation et de financement sur la production des savoirs et les contenus scientifiques (Barrier, 2011; Jouvenet, 2011; Braun, 1998; Gläser, Laudel et Whitley, 2018; Gläser et Laudel, 2016; Laudel, 2017). Enfin, certaines enquêtes sont consacrées à la manière dont les universitaires s'approprient des classements, des indicateurs et des standards, en montrant comment ils composent avec ces dispositifs ou développent des stratégies pour y résister (Hallett, 2010; Espeland et Sauder, 2007; Paradeise et Thoenig, 2013; Hazelkorn, 2015).

En multipliant les points de vue sur les secteurs professionnels et sur les pays (Ferlie, Musselin et Andresani, 2009), ces enquêtes ont alimenté un riche débat quant au caractère plus ou moins structurant des réformes et des outils de pilotage contemporains sur le travail, les identités et les pratiques des professionnels. Cependant, ceux-ci sont souvent dépeints en situation de subir les restructurations de leur travail et de s'en accommoder. Centrées sur les effets des réformes sur les pratiques des professionnels, les enquêtes laissent souvent dans l'ombre l'influence et le rôle éventuels de certains d'entre eux dans l'élaboration des instruments managériaux et de politiques visant à définir et normer le travail par des procédures¹⁴. En s'intéressant à la fabrique des dispositifs d'évaluation scientifique, l'ouvrage entend adopter un angle original, dépassant l'alternative entre maintien ou perte du pouvoir professionnel. Il montrera plutôt comment se reconfigure le pouvoir scientifique – autrement dit, comment et pourquoi certains membres de la profession endossent des discours managériaux et comment ils réinventent les définitions de la valeur scientifique.

ENQUÊTER SUR LA FABRIQUE DU JUGEMENT SCIENTIFIQUE

Dans le secteur de la recherche, d'importants travaux ont mis en lumière les transformations contemporaines des pratiques des chercheurs et de l'organisation du travail. La boîte noire de l'évaluation et des dispositifs qui orchestrent la répartition du crédit académique est cependant plus

14 À de notables exceptions : en Grande-Bretagne, les travaux de Maurice Kogan et Steve Hanney (2000) montrent par exemple qu'une fraction d'universitaires britanniques a été investie dans la création et la mise en place du RAE.

rarement ouverte¹⁵. Cet ouvrage propose de renverser la perspective, en étudiant des scientifiques non pas en train de « faire avec » des réformes dans leur quotidien de travail, mais en train de bâtir et de débattre des normes d'évaluation régulant le travail de leurs collègues.

Cette approche ouvre de nouvelles interrogations. Comment s'est construit, en dépit des critiques qu'il suscite, ce dispositif d'évaluation de la recherche scientifique ? Qui sont les acteurs qui le portent ? Quels savoirs mobilisent-ils pour réformer les règles de la reconnaissance scientifique ? Comment les normes d'évaluation qu'ils inventent ou réinventent s'articulent-elles avec des pratiques épistémiques propres aux disciplines ou les concurrencent-elles ? Dans quelle mesure, enfin, redéfinissent-elles la valeur de la science et ce que « faire de la recherche » veut dire aujourd'hui ?

En entrant par la petite porte d'une agence contestée, cet ouvrage offre un éclairage inédit sur les transformations du gouvernement du monde scientifique au tournant du XXI^e siècle. Au plus près des acteurs qui font et défont les règles de la reconnaissance scientifique, il propose de comprendre comment des acteurs du monde scientifique se mobilisent pour promouvoir des réformes et construire un nouveau dispositif de régulation professionnel. Il montre ainsi que les débats sur les pratiques de recherche légitimes et les bonnes manières d'en juger n'opposent pas des scientifiques et des *managers*, mais qu'ils se nichent au sein même de la profession scientifique. L'enquête met alors en lumière l'intrication permanente entre savoirs et action publique, en montrant que la construction ou déconstruction des instruments d'évaluation est le produit de débats épistémiques sur la science, mais aussi de rapports de pouvoir entre les scientifiques qui peuplent les instances du gouvernement de la profession.

Savoirs, disciplines et action publique

Une place centrale est accordée dans cet ouvrage aux conceptions de la recherche qui informent les instruments d'action publique, et réciproquement, aux façons dont ces instruments redessinent les activités scientifiques légitimes. En effet, les producteurs des dispositifs d'évaluation arriment leur activité de production de normes à des représentations du

15 À de rares exceptions, comme les travaux de Michèle Lamont (2009) qui entre dans les coulisses de l'évaluation des projets de recherche aux États-Unis, ou ceux de Christine Musselin (2005) sur les pratiques d'évaluation et de recrutement universitaires.

fait scientifique, et à leurs propres pratiques d'organisation et de production du travail académique. La question des savoirs qui informent l'action publique se pose avec acuité dans une agence qui centralise et uniformise les procédures de jugement : les pairs ne sont pas tous égaux au moment d'énoncer les critères qui comptent. Ce sont par exemple des physiciens et des médecins qui ont occupé des positions centrales dans les premières années d'existence de cette agence. Il faut attendre quatre ans après la mise en place de l'AERES pour qu'un littéraire accède à une position de direction en matière d'évaluation de la recherche. Il n'est pas surprenant que des académiques issus des SHS aient pris la parole dans les espaces médiatique et scientifique pour exprimer l'inadéquation des méthodes d'évaluation de l'AERES, dénoncer l'alignement des standards d'évaluation sur ceux des sciences expérimentales, ou s'ériger contre le « diktat » d'une évaluation quantifiée. Des numéros de revue (Zarka et Barbéris, 2009) et des travaux collectifs (Cassin, 2014) procèdent à l'analyse critique des instruments de mesure promus au sein de cette agence, et à l'idéologie générale dans laquelle prennent place ces réformes.

L'ouvrage questionne dans quelle mesure et comment les instruments retenus reflètent des conceptions de la science et des normes disciplinaires dominantes. Ce faisant, on sera attentif à la vie mouvante de ces instruments : faiblement stabilisés, ils sont réinventés au cours de l'existence de l'AERES, si bien que les rapports de domination apparaissent fluctuants. Ainsi, certains instruments ont pu traduire l'assujettissement des certaines disciplines des SHS à des modèles d'évaluation et des pratiques de travail qui leur sont étrangers. Mais les représentants de ces disciplines à l'AERES sont parvenus à intervenir dans la définition des critères qui comptent pour eux : ils ont eu carte blanche pour les réformer au cours d'une consultation de plusieurs mois en 2011, à laquelle j'ai assisté en tant qu'observatrice. Il s'agira de comprendre comment chercheurs et universitaires se mobilisent, avec plus ou moins de succès, pour faire reconnaître dans les instruments d'évaluation les définitions de la qualité académique telles qu'ils les conçoivent et les projets de politiques scientifiques qu'ils défendent.

Les clivages internes au monde scientifique se donnent particulièrement à voir dans les débats virulents autour de certains instruments d'évaluation. Tel est le cas des classements de revues¹⁶. Alors que certains littéraires à l'AERES ont milité pour que soit introduit un tel classement

16 Le chapitre 3 revient en détail sur les controverses liées à cet instrument indirect d'évaluation de la qualité des publications scientifiques.

dans leur domaine, leurs collègues et représentants d'autres instances de la discipline s'y sont massivement opposés. Des membres du Conseil national des universités (CNU)¹⁷, du CoNRS et des associations savantes se sont constitués en contre-pouvoirs efficaces : ils ont fait valoir des discours sur les caractéristiques épistémiques de la littérature pour entraver la carrière d'instruments auxquels ils n'accordent pas crédit.

Les évolutions de l'action publique peuvent aussi ouvrir de nouveaux conflits internes à la communauté académique, voire les redessiner à l'intérieur même des disciplines. En effet, l'agence étudiée n'est pas un isolat social, et les académiques peuvent également peser, de l'extérieur, sur l'élaboration et sur l'architecture des instruments. Ainsi, si la production des instruments est le produit de conflits sur la définition de la science, elle s'encastre également dans des rapports de pouvoir plus contingents et évolutifs entre des représentants d'institutions concurrentes du gouvernement de la science.

Enfin, l'action publique est façonnée par des savoirs et des représentations du fait scientifique, mais ce processus est aussi réciproque : les instruments retenus, en encadrant la production du jugement, ont aussi des incidences sur les définitions de la science et des pratiques académiques légitimes. L'ouvrage interroge les façons dont les chercheurs s'approprient ces instruments dans leurs pratiques de jugement. Suivre les scientifiques au travail, dans leurs activités d'évaluation, montre que cette agence n'uniformise pas les modalités de la reconnaissance professionnelle. Celles-ci varient en effet considérablement selon les disciplines. Cependant, les procédures et les critères de l'AERES redéfinissent les conditions auxquelles doivent souscrire travaux de recherche et organisations scientifiques pour être valorisées. Plus encore, les scientifiques peuvent s'appuyer sur ces instruments pour accompagner des transformations qu'ils estiment bénéfiques, concernant les pratiques d'organisation du travail scientifique et les méthodes de validation du savoir, alors même que ces réformes ne font pas toujours consensus au sein de leur discipline. Ce livre ouvre une réflexion sur les espaces inattendus où ces instruments agissent, en transformant les conceptions de la qualité et du travail scientifiques.

17 Le CNU est une instance nationale qui se prononce sur les mesures individuelles relatives à la qualification, au recrutement et à la carrière des universitaires.

Pour étudier ces controverses internes au monde scientifique, l'ouvrage s'appuie sur une enquête de terrain combinant entretiens approfondis, sources écrites – archives personnelles et nationales, pamphlets et pensées libres recueillis sur des blogs publics, des sites d'associations savantes et de syndicats de la profession – et observations directes de commissions où sont construits les critères de jugement. Par ailleurs, le caractère à la fois public et prolifique des productions écrites de l'AERES les prête à des traitements quantifiés. À l'appui de l'analyse qualitative, la construction et l'exploitation de deux bases de données complètent l'éventail des méthodes employées¹⁸. Combiner ces sources et ces méthodes d'analyse permet de jouer avec différentes échelles d'analyse : les commissions observées sont des dispositifs privilégiés pour étudier dans le détail, le plus finement possible, les enquêtés au travail, tandis que l'analyse quantitative s'avère utile pour resituer les cas ethnographiques dans un ensemble de pratiques plus vaste qui les englobe. Les méthodes utilisées sont classiques en sociologie, mais deux particularités de l'objet d'étude invitent à une réflexion spécifique : travailler sur une institution contestée au sein du monde scientifique, et étudier un milieu professionnel dans lequel j'étais moi-même immergée et partie prenante, sans être pour autant l'égale de mes « pairs », appellent un retour sur la gestion de la proximité mais aussi de la distance au « métier ».

La première fois que je suis entrée dans les locaux de l'AERES, rue Vivienne à Paris, presque encore neufs et confortablement aménagés, c'était en juin 2011, soit quelques mois avant le début de la thèse en sociologie dont est issu cet ouvrage. À cette époque, l'AERES avait organisé un colloque sur les indicateurs d'évaluation en SHS, auquel j'ai assisté. M'intéressant aux pratiques de jugement dans ces disciplines dans le cadre de mon doctorat, j'ai manifesté mon intérêt pour les activités conduites dans l'agence sur ce sujet. C'est ainsi que j'ai appris la création d'une commission de réflexion interne, échelonnée sur plusieurs mois à partir de septembre 2011, pour refonder les critères d'évaluation de la recherche en SHS. Des membres de la commission m'ont conviée à y assister en tant qu'observatrice après avoir signé un accord de confidentialité,

18 Pour plus de précisions sur le dispositif d'enquête et d'analyse, voir l'annexe 2, « Sources, méthodes et principes d'anonymat ».

qui m'engageait seulement à exploiter les données recueillies dans le cadre de travaux et de publications académiques (et non journalistiques).

J'ai été dans un premier temps surprise par la facilité d'accès au terrain à l'AERES, étant donné les tensions que sa création avait soulevées dans le milieu académique. En effet, non seulement j'ai été accueillie au sein de cette commission de réflexion interne, à laquelle étaient aussi invités deux autres chercheurs en sciences sociales à titre d'observateurs, mais encore presque tous les entretiens sollicités auprès des scientifiques et des administratifs¹⁹ de l'AERES ont été acceptés, parfois avec un certain enthousiasme. Si leur ouverture à un regard extérieur sur leurs pratiques semble coïncider avec l'apologie de la « transparence » de l'évaluation, mise en avant par une majorité des membres de l'agence, il m'est apparu que c'était également parce que l'agence était controversée que ses membres ont reçu avec bienveillance et intérêt mes demandes de rendez-vous. Plusieurs d'entre eux, en fonction au moment de l'enquête, considéraient qu'une étude empiriquement fondée et en quelque sorte dépassionnée sur leur activité rendrait justice au travail d'évaluation à l'AERES tel qu'ils le conçoivent. Ou, comme me le confiait l'un d'entre eux : « Montrez-leur que nous ne sommes pas des voyous ! ». Ainsi, bien que je n'aie été tenue d'aucune sorte à leur soumettre mes travaux avant publication, mon enquête était, aux yeux de certains d'entre eux, susceptible de conforter la crédibilité publique de leur activité normative, et notamment face aux détracteurs de l'agence.

Le fait de me présenter comme doctorante, interviewant ou observant des professeurs – majoritairement des hommes – installés dans la carrière, a pu conduire à des situations asymétriques. En entretien, il arrivait que les enquêtés commencent par me poser des questions, avant que je puisse prendre la main sur la conduite de l'entretien. Certains m'ont par exemple interrogée sur mon parcours, les résultats de mes travaux, ma démarche de recherche, voire sur mon plan de thèse. Cette inversion est fréquente dans les situations d'entretien (Chamboredon *et al.*, 1994), mais elle prenait un sens spécifique ici : elle ressemblait parfois à s'y méprendre à une

19 La distinction entre « administratifs » et « scientifiques » se réfère à ce qui constitue l'activité professionnelle principale pour laquelle ils sont rémunérés. Les scientifiques qui composent l'AERES occupent ou ont pu occuper des fonctions administratives (c'est d'ailleurs en partie ce qui les distingue dans la communauté académique), mais leur cœur de métier, qu'ils revendiquent souvent dans les entretiens, est l'enseignement et la recherche. Le terme « scientifique » ou l'anglicisme « académique » désigne dans cet ouvrage de façon générique ces professionnels, quels que soient leur statut et leur discipline d'appartenance.

épreuve de validation (scientifique) de mon travail par des chercheurs. Même si je ne m'inscrivais pas dans leur discipline, certains me prodiguaient des conseils de nature variée, d'ordre méthodologique (« J'essaye de voir si vous aviez appris à faire des entretiens » (professeur, STAPS), ou théorique (« À mon avis, vous devriez réfléchir à [telle] question de recherche » (professeur, épistémologie). D'autres mettaient davantage à l'épreuve ma position dans le monde académique, interrogeant mes préférences : « Vous avez des options en sociologie ? Des auteurs particuliers, préférés, fétiches ? Quel est votre *top five* ? » (professeur, philosophie), et jugeant ainsi ma compétence.

Le fait de présenter ces extraits hors contexte ne doit pas conduire à des mésinterprétations : de très rares exceptions, ces interventions relevaient plutôt d'une prise de distance avec la situation d'entretien même, avec laquelle certains n'étaient pas nécessairement à l'aise. Bien souvent bienveillante, leur attitude s'enracinait généralement dans la curiosité qu'ils manifestaient pour un sujet d'investigation qu'ils trouvaient quelque peu étrange (l'évaluation de la recherche) et qui les concernait directement. C'est pourquoi je parvenais rarement à m'éclipser sans avoir répondu à ces questions relatives à mon enquête, à la fin de l'entretien si ce n'était au début.

J'ai donc eu assez rarement l'impression d'affronter des « imposants » (Chamboredon *et al.*, 1994). Lorsque les rapports de domination symbolique étaient manifestes, cette situation pouvait être inconfortable sans être incommode pour la qualité de l'entretien : l'argumentation structurée et détaillée que me proposaient ces acteurs, maîtres face à l'élève, le soin avec lequel ils clarifiaient leur discours à la suite de mes questions, m'a permis de recueillir des informations et un matériau d'enquête fournis. Par ailleurs, une part bien plus importante des enquêtés me renvoyait en entretien l'image d'une « future collègue » à même de comprendre leurs conditions de travail et ses enjeux : ils m'incluaient volontiers dans la confiance, me faisant assez spontanément confiance pour ne pas trahir leurs propos et savoir les écouter en connaissance de cause. La question qui se posait était plutôt celle de la gestion des différentes formes de proximité et de distance avec les enquêtés. Ayant un pied dans le métier, en tant qu'« apprentie », j'ai pu être perçue par les professeurs comme plus proche d'eux que des journalistes, par exemple, qui viendraient enquêter sur eux sans appartenir au monde académique. Mais les représentants de certaines disciplines m'ont parfois renvoyé l'image d'une étrangère – ce qui a pu entraver le cours de l'échange.

Tel a été le cas lors d'un entretien avec une professeure émérite en littérature, qui était fortement réticente lors de notre échange et répondait à chaque question, si ouverte soit elle, en quelques mots. Si cet échec à obtenir la parole de l'enquêtée peut être lié au simple fait que cette enquête ne l'intéressait pas, il y a pu également avoir mécompréhension quant à la nature de l'entretien. Le fait d'avoir précisé que je réalisais une thèse en sociologie a pu la conduire à répondre à mes questions comme s'il s'agissait d'un questionnaire voire d'un sondage – méthode qu'elle était susceptible de considérer comme typique des sciences sociales –, et ce d'autant plus que j'ai été amenée, pour des raisons matérielles, à réaliser cet entretien par téléphone²⁰. Dans ces cas, c'était à mon tour de construire une intimité et une empathie avec les enquêtés. Mon parcours académique a ici été une ressource précieuse pour construire une telle proximité, à la fois disciplinaire et institutionnelle : avant d'entreprendre une thèse en sociologie, j'étais élève à l'École normale supérieure de Lyon en littérature, où ont été formés de nombreux enquêtés dans cette discipline. Lors d'un entretien, un évaluateur s'enquêrait, un peu soucieux, du regard que je pouvais porter sur ses pratiques, me demandant si je considérais les littéraires comme des « animaux bizarres ».

Avec des professeurs appartenant à des disciplines littéraires, le fait de signaler que j'avais été formée en littérature et en langues permettait souvent d'estomper d'éventuelles incertitudes quant à ma position d'enquêtrice et d'observatrice extérieure, en plaçant l'échange sur le registre de la connivence. Considérant que je « connai[sais] la maison », pour reprendre les termes de certains enquêtés, ils m'ont livré plus aisément leur interprétation de ce que signifiait faire de la recherche dans ces disciplines, et des évolutions qu'elles connaissent dans un contexte de réformes de l'évaluation, sur la base d'un intérêt mutuel pour ces questions.

La question du rapport au terrain, qui se pose au moment de la réalisation de l'enquête, se retrouve aussi une fois celle-ci achevée, au moment de l'analyse. Le fait de combiner les données, de croiser les sources, m'a aidée à trouver la juste distance avec l'enquête et ses acteurs : ni mystifié, ni désabusé ou suspicieux quant à leurs discours et leurs pratiques, cet ouvrage s'efforce de les contextualiser et d'en comprendre les logiques.

20 Seuls quatre entretiens ont été réalisés par téléphone. La bonne volonté des enquêtés a généralement permis de surmonter les difficultés liées à l'absence d'interaction en face à face, mais nous n'y avons eu recours que lorsqu'il n'était pas possible de nous rencontrer *de visu*.

Parcourant les coulisses de l'évaluation scientifique, la réflexion se centre sur les scènes majeures où sont produits, débattus, appropriés les instruments d'évaluation scientifique depuis la création de l'AERES. Je fais de ces arènes encore peu explorées des observatoires privilégiés de la façon dont la profession académique définit le fait scientifique et les manières légitimes d'en juger.

Le premier chapitre montre que la mise en place du dispositif d'évaluation à l'AERES n'est pas le produit d'un consensus réformateur. Elle est au contraire l'enjeu de rapports de force entre différents groupes professionnels et entre différentes institutions du gouvernement de la recherche. M'intéressant aux acteurs qui parviennent, à l'issue de ces luttes, à capter le pouvoir de définir les instruments d'évaluation de la recherche, je montre qu'il ne s'agit pas d'experts situés hors de la profession académique, mais au contraire, de scientifiques en exercice. Ce chapitre en propose, par l'analyse biographique, un portrait détaillé.

Que font donc ces scientifiques à l'agence, et pourquoi s'investissent-ils dans cette nouvelle institution controversée ? Comment leurs convictions se traduisent-elles dans des instruments d'action publique ? Le deuxième chapitre montre comment les membres de l'agence réinvestissent leur parcours et leur expérience dans les institutions du monde académique pour défendre la nécessité de mettre en règle la production du jugement. Il éclaire le paradoxe selon lequel des académiques mobilisent une rhétorique « professionnelle » pour bureaucratiser la fabrique du jugement et promouvoir un contrôle accru des pratiques de leurs collègues.

Cependant, ce n'est pas parce que les membres de l'AERES définissent des « bonnes pratiques » et des instruments d'évaluation que ceux-ci s'imposent dans le monde scientifique. Certains d'entre eux subissent au contraire de puissantes remises en cause, et la critique permet, à certaines conditions, de déconstruire l'action publique. C'est à ces processus que s'intéresse le troisième chapitre, qui prend le cas de la carrière avortée des classements de revues en littérature. Questionnant les ressorts de la mobilisation réussie des littéraires contre l'instrument, il décrypte les stratégies et les argumentaires épistémiques déployés par les représentants de la discipline pour tenir en échec cette réforme.

Les instruments d'évaluation de cette agence ne sont pas immuables dans le temps. Ils se transforment au gré des contestations qu'ils génèrent au sein de la communauté scientifique.

À partir de l'observation directe d'une consultation organisée par les représentants des SHS à l'AERES dès 2011, le quatrième chapitre étudie le travail conduit par ses membres pour réinventer le dispositif d'évaluation scientifique. Il met en lumière les savoirs qu'ils mobilisent pour construire un nouvel instrument d'appréciation de la valeur des travaux de leurs collègues, et les enjeux des controverses qui les mettent aux prises. Les membres de la commission revêtent alors un rôle de sociologue des sciences, en établissant, avec prudence, les frontières de l'activité académique : c'est finalement le travail de définition de la science légitime en train de se faire, dans ses incertitudes et ses tâtonnements, que donne à voir ce chapitre.

Enfin, dans le cinquième chapitre, je m'efforce de comprendre comment les universitaires et les chercheurs s'approprient ces instruments lorsqu'ils évaluent les travaux de leurs collègues. Pour cela, j'ai choisi de comparer les pratiques de jugement dans deux disciplines des sciences humaines et sociales, la littérature et la géographie, qui diffèrent par leur histoire et les pratiques de recherche qui s'y sont développées. La comparaison montre que certains universitaires discréditent et contournent ces instruments dans la production de leur jugement, tandis que d'autres en font un usage stratégique. Bien que les standards de l'AERES pèsent inégalement sur les pratiques de jugement, ils contribuent néanmoins à rendre visibles certaines conceptions de la recherche et des « bonnes manières » de produire des savoirs.