

De l'évaluation collégiale à l'évaluation à dominante gestionnaire

Le cas de l'AERES

Robert Boure



Édition électronique

URL : <http://journals.openedition.org/communicationorganisation/1391>

DOI : [10.4000/communicationorganisation.1391](https://doi.org/10.4000/communicationorganisation.1391)

ISSN : 1775-3546

Éditeur

Presses universitaires de Bordeaux

Édition imprimée

Date de publication : 1 décembre 2010

Pagination : 39-62

ISBN : 978-2-86781-743-4

ISSN : 1168-5549

Référence électronique

Robert Boure, « De l'évaluation collégiale à l'évaluation à dominante gestionnaire », *Communication et organisation* [En ligne], 38 | 2010, mis en ligne le 01 décembre 2013, consulté le 03 mai 2019. URL : <http://journals.openedition.org/communicationorganisation/1391> ; DOI : [10.4000/communicationorganisation.1391](https://doi.org/10.4000/communicationorganisation.1391)

De l'évaluation collégiale à l'évaluation à dominante gestionnaire

Le cas de l'AERES

Robert BOURÉ⁰¹

Comme d'autres termes qui font désormais partie du vocabulaire courant des SHS, celui d'« évaluation » est tiraillé entre des significations situées à des niveaux différents. En effet, il désigne à la fois une notion travaillée par les SHS et la philosophie, des pratiques sociales et des dispositifs. En outre, comme pour tous les mots affectés par la banalisation de leurs usages scientifiques et pratiques, chaque niveau de signification demande à être contextualisé dans l'espace, le temps et les différents champs concernés par ces usages.

En fait, le terme a une étrange destinée. Alors qu'il est présent dans les anciens dictionnaires de la langue française (*Dictionnaire universel de Furetière*, *Robert historique de la langue française...*) où il renvoie à l'estimation de la valeur d'un objet au sens soit de mesure (d'un prix, d'une quantité, d'une distance), soit de jugement de valeur, il est longtemps largement ignoré par la plupart des dictionnaires de SHS⁰², preuve supplémentaire de l'historicité des

01. Professeur de Sciences de l'information et de la communication, chercheur au LERASS, équipe Médiapolis, Université de Toulouse. Cet article s'appuie moins sur des travaux scientifiques publiés, encore peu nombreux, que sur des débats (notamment ceux organisés à l'EHESS en 2008-2010) et prises de position mis en ligne ou pointés avec un lien sur des sites comme Sauvons l'université, Sauvons la recherche, Évaluation de la recherche en SHS. Il s'agit majoritairement de textes : 1) critiques, pour ne pas dire hostiles ; 2) rédigés plus par des chercheurs que par des enseignants-chercheurs, de sorte que l'enseignement est peu abordé ; 3) publiés avant que l'AERES atteigne sa vitesse de croisière. Il s'appuie aussi sur des documents produits par l'AERES ainsi que sur des observations personnelles liées aux expertises réalisées pour le compte de l'AERES (comités de visite, rédaction de rapports, participation à des réunions de statut divers...) et aux situations d'évalué par cette instance (co-responsable d'une équipe dans une EA et responsable de deux diplômes).

02. On notera cependant que dans quelques disciplines, la notion a été rapidement construite comme un enjeu scientifique et pratique, une des plus emblématiques étant les sciences de l'éducation où son usage se répand dès leur reconnaissance officielle, dans les années 1970. Pour sa part, la philosophie a intégré depuis longtemps le terme à son vocabulaire, et notamment au niveau de la morale (elle renvoie aux jugements de valeur) et de l'épistémologie générale où elle est une procédure destinée à déterminer la valeur épistémique d'une théorie, autrement dit sa capacité à produire des connaissances, ou à comparer la valeur épistémique de plusieurs théories.

objets de ces disciplines. Mais à partir des années 1990, il s'installe rapidement dans ces usuels, la pratique linguistique savante semblant alors s'aligner sur les pratiques sociales et les dispositifs d'évaluation profanes. En effet, ces derniers sont désormais omniprésents dans de nombreux domaines (aménagement, santé, action sociale, diplomatie, police, culture, enseignement, recherche...) et pour toutes les catégories d'acteurs (collectivités publiques, entreprises, associations...), de sorte que l'on est fondé à parler d'évaluation généralisée et, avec Valérie Péresse, à ériger l'évaluation en « impératif absolu ».

En même temps, l'évaluation s'est profondément transformée puisqu'elle est désormais présentée comme un fondement important de la prise de décision et plus généralement de l'action. L'action est interrogée par elle à partir de critères quantitatifs et qualitatifs se voulant « objectifs » et sous plusieurs angles complémentaires : pertinence (par exemple, par rapport aux attentes des acteurs), effectivité, efficacité (les objectifs sont-ils atteints ? quels sont les effets positifs et négatifs directs ou induits ?), efficience (au regard des moyens, et en particulier des moyens financiers), cohérence (adéquation des objectifs et des moyens), viabilité et continuation. Vis-à-vis de l'action, elle est devenue un outil de réflexion, de gestion, de régulation et enfin de légitimation puisqu'elle la justifie, d'une part, en prétendant s'appuyer sur des dispositifs rigoureux et méthodiques, d'autre part, en faisant d'elle un attribut indispensable de la modernisation des appareils publics et privés, voire quand il s'agit des collectivités publiques, un attribut de la démocratie *via* la transparence.

Elle ne saurait donc être réduite à des finalités pratiques. Selon Georges Balandier (2010, 9-10), l'évaluation a une forte dimension symbolique et plus précisément idéologique dans une période où « les idéologies de confrontation rejoignent les langages perdus » et où les sociétés vivent dans l'immédiat. L'évaluation apparaît alors comme une idéologie de substitution qu'il place au même niveau que la croyance religieuse, « deux figures symétriques qui disent répondre aux incertitudes (généralisées) et aux impuissances du présent ». Pour Albert Ogien, il s'agit moins de la violence symbolique d'une idéologie ou de la violence institutionnelle d'un pouvoir que de la violence arithmétique de la quantification. En même temps, il fait sinon de l'évaluation, du moins de ses formes contemporaines, un « principe politique »⁰³.

Dans l'enseignement supérieur et la recherche français, la généralisation de l'évaluation et l'évolution rapide de ses formes⁰⁴ revêtent une signification particulière, notamment en raison du rôle traditionnellement attribué à l'évaluation par l'ainsi nommée « communauté scientifique »⁰⁵. Alors

03. Voir aussi, *Cités*, 2009 et *Cahiers Internationaux de Sociologie*, 2010.

04. Cette évolution et les problèmes qu'elle pose aux évalués (institutions académiques et individus) ne sont pas propres à la France (Servais, 2011).

05. Nous n'entendons pas ici discuter la pertinence de cette expression. Pour une présentation très « pédagogique » des débats auxquels elle donne lieu depuis l'entre-deux guerres, cf. Dubois, 1999.

que l'évaluation en tant que principe et pratique communautaires est une constante de la vie académique, nombre d'enseignants-chercheurs et de chercheurs accueillent avec méfiance, quand ce n'est pas rejet, celle qui leur est aujourd'hui appliquée et qui s'incarne dans des structures nouvelles créées pour être pérennes (Agence d'évaluation de la recherche et de l'enseignement supérieur -AERES- et Agence nationale de la recherche -ANR-) ou, à un degré moindre, dans des instances *ad hoc* comme les comités de sélection. Ce qui fait l'objet des critiques les plus vives, c'est moins la contestation de tel ou tel élément des dispositifs que le cadre général en fonction duquel ces derniers ont été conçus. Autrement dit, « un régime d'évaluation inédit (...) qui tend aujourd'hui à s'imposer » (Dodier, 2009, 2) et qu'Albert Ogien (2009, 1) nomme « évaluation gestionnaire », formule abrupte à laquelle on substituera celle d'« évaluation à dominante gestionnaire » car ce régime contient encore des éléments relevant de modes antérieurs.

D'un régime d'évaluation à un autre

L'évaluation collégiale

Qu'on la nomme évaluation « collégiale », « par les pairs » ou plus simplement « professionnelle », elle présente les caractéristiques suivantes

- Inscription dans une double tradition. D'une part, une tradition ancienne, schématiquement celle de l'auto-organisation et de l'autocontrôle de la communauté scientifique qui s'affirme dès le XVII^e siècle avec l'avènement de la science moderne et de certains de ses corollaires institutionnels tels les Académies scientifiques, ou symboliques comme la République des Lettres⁰⁶; dans cette perspective, l'évaluation collégiale est une « *manière de faire exister le sens de la communauté scientifique dans les dispositifs d'évaluation* » (Dodier, 2009, 12). D'autre part, une tradition plus récente, la protection et la réglementation par un État républicain d'autant plus « bienveillant et respectueux » (bien que souvent « radin ») vis-à-vis de l'enseignement supérieur et de la recherche qu'il a érigé en services publics; on peut d'ailleurs voir dans l'évaluation collégiale française une manière de concilier deux registres de l'*ethos*, celui de la science (cf. par exemple Merton, 1973), et celui du service public (souci de l'égalité et de la démocratisation, dévouement, vocation...).

- Collégialité : le principe de base veut que les pairs se réunissent entre eux dans des comités et commissions pour élaborer des règles de fonctionnement et des critères d'appréciation. Il s'agit *in fine* de construire un tiers partagé interne, c'est-à-dire des référents communs relativement stables, publics et acceptables par le milieu professionnel auquel ils appartiennent, et au-delà par

06. Cette expression désigne la réunion symbolique des lettrés, érudits ou savants, termes alors aux significations voisines. Elle renvoie à l'idée que ses membres appartiennent à une sphère ayant des valeurs, des manières de faire et des fonctionnements spécifiques (salons, voyages, échanges réguliers de correspondances, Académies...), distincts des autres métiers.

la communauté scientifique dans son ensemble. Ces membres des commissions ne sont pas des permanents (ils ont un mandat formel ou informel), ils disposent du temps nécessaire à l'accomplissement de leurs tâches et ils sont amenés à se revoir régulièrement, ces deux dernières conditions excluant ou relativisant fortement les instances *ad hoc*. Enfin, chacun est placé sous le regard des autres commissionnaires, voire sous celui de la communauté.

- Autonomie relative des instances d'évaluation, découlant de la relative autonomie de l'activité scientifique: elle est garantie par l'État et elle s'exprime notamment par la systématisation des procédés de la cooptation et de l'élection (ou de la présence d'élus), même si les tutelles ont souvent cherché à injecter des doses de nominations.

- Finalité ultime « globalement » partagée par les évaluateurs: favoriser la production ou la transmission des connaissances, ce qui revient à évaluer la contribution que les individus et les institutions apportent à leur construction et à leur diffusion. Cela n'exclut pas d'autres finalités répondant à des logiques particulières (institutionnelles, disciplinaires, personnelles, réseautiques -cf. les collèges invisibles, au sens de Solla Price, 1972-, de « champ » -cf. Bourdieu, 1975, 1984-), mais les relativise.

- Pluralisme: la collégialité, l'autonomie et le consensus sur la finalité principale font de chaque membre non seulement un représentant d'un milieu, d'une manière de « faire science » ou de « faire enseignement », mais aussi quelqu'un dont on attend qu'il comprenne qu'il y a d'autres milieux que le sien et d'autres manières de faire que les siennes (Dodier, 2009). La négociation est *de facto* au poste de commande, de sorte que les décisions stratégiques sont souvent le fruit de compromis négociés et d'ajustements successifs. Cela revient à accepter collectivement non seulement l'existence de plusieurs manières de « faire science » (paradigmes, méthodes, objets...) ou de « faire enseignement », mais encore que l'évaluation d'un objet (activité d'un laboratoire, valeur d'une formation, recrutement ou promotion...) ne présuppose pas un accord complet sur le fond.

Si aucune instance d'évaluation collégiale ne fonctionne ou n'a fonctionné suivant le modèle idéal aux accents quelque peu mertonien, on peut en relever plusieurs qui s'inspirent plus ou moins largement de sa « philosophie générale »: comités de lecture de revues, comités scientifiques de colloques et congrès, sections du Conseil national des universités (CNU), comités et commissions des organismes de recherche (CNRS, INSERM, INRA...), commissions de spécialistes, Conseils scientifiques et Conseils des études et de la vie universitaire des universités...

Pourtant le modèle et les instances qui s'en inspirent ont fait l'objet de multiples critiques. On laissera définitivement de côté celles que la sociologie des sciences post-mertonienne a adressées à un mode d'organisation et à des valeurs un peu trop idéalisées et provisoirement celles que les enseignants-chercheurs et les chercheurs ont élaborées, parfois de façon impressionniste,

pour se concentrer sur les critiques venant des sphères gouvernementales. Si l'on relativise celles qui se situent dans le registre de la provocation (cf. par exemple le discours de Nicolas Sarkozy du 22 janvier 2009⁰⁷) et qui dénoncent pêle-mêle les crispations corporatistes, l'immobilisme, l'intérêt porté plus aux moyens qu'aux résultats, quand ce n'est pas le rôle néfaste des syndicats, il reste deux critiques majeures, d'ailleurs déjà formulées dans d'autres domaines. Les formes d'évaluation traditionnelles seraient :

- Des freins à « l'indispensable » modernisation de l'Administration en général et à celle de l'enseignement supérieur et de la recherche en particulier. Il faut donc définir un régime d'évaluation adapté à la restriction du périmètre d'intervention d'un État devenu stratège (Bezes, 2005), à l'évolution vers la « gouvernance » (*Sciences de la Société*, 2003 ; Crowley, 2004 ; Bevir, 2010), mais aussi aux exigences de la Loi Organique relative aux Finances Publiques (LOLF), de la Révision Générale des Politiques Publiques (RGPP)⁰⁸, de la LRU, de la énième réforme du CNRS, de la création des pôles d'enseignement supérieur et de recherche (PRES)...

- En décalage avec les exigences de la « nécessaire » construction de l'espace européen de l'enseignement supérieur et plus généralement de la tout aussi « nécessaire » compétition internationale (Lesresche, Larédo, Weber, 2009), symbolisée par le classement de Shanghai.

Ces critiques tirent leur force autant de la force de l'évidence que de celle de ceux qui les portent, les nouveaux gestionnaires publics qui « exercent un pouvoir de conviction et de ralliement à tous les échelons de l'appareil étatique » (Bruno, 2008, 35). En tout état de cause, elles débouchent sur des exigences nouvelles. Et à exigences nouvelles, régime d'évaluation nouveau.

L'évaluation à dominante gestionnaire

Cette évaluation est très différente dans sa philosophie comme dans ses dispositifs techniques des modes d'évaluation dont le monde de l'enseignement supérieur et de la recherche s'était progressivement doté.

07. A l'occasion du lancement de la réflexion pour une stratégie nationale de recherche, le Président de la République prononce à l'Élysée un discours au cours duquel, entre autres, il dénonce le mode d'évaluation par les pairs en des termes souvent jugés provocateurs. Ces propos entraînent des protestations nombreuses et vigoureuses : pétition pour exiger des excuses (plus de 13 342 signataires dont 10 000 en quelques jours), réactions publiques de scientifiques reconnus, prises de positions d'instances (dont l'Académie des sciences), lettre de 122 membres de l'Institut universitaire de France (11 février 2009).

08. La LOLF centre la procédure budgétaire autour de la performance de la gestion publique et de la transparence de l'information budgétaire. L'enseignement supérieur et la recherche sont concernés : ainsi, par exemple, la Mission pour la recherche et l'enseignement supérieur du Ministère contient plusieurs « projets annuels de performances » ou « PAP » (voir www.performance-publique.gouv.fr) définissant des « cibles de performances ». La RGPP est une remise à plat des politiques publiques qui s'attache principalement à la réduction des dépenses publiques. Si la RGPP ne concerne pas pour l'instant l'Université, elle va toucher en 2012 le CNRS, l'INSERM, l'INRA, l'INRIA et le CIRAD sur les fonctions autres que la recherche, en tout cas pour l'instant.

1) Elle est largement fondée sur le *New public management*⁰⁹ et plus précisément sur le contrôle de la performance par l'appréciation du résultat à travers une démarche se voulant objective dans ses conditions d'exercice (transparence, neutralité, technicité, professionnalisme), ses critères quantitatifs et qualitatifs et ses outils dont la dénomination trahit les origines géographiques, le *benchmarking* (ou comparaison par la performance qui permet les classements) étant le plus emblématique (Bruno, 2008). En même temps, elle met très fortement l'accent sur la responsabilité, la qualité, l'apprentissage organisationnel et, bien entendu, l'efficacité. Et c'est bien souvent davantage l'efficacité des dépenses publiques que l'on cherche à apprécier que celle des missions dont ces dépenses doivent garantir l'exercice, notamment en termes de moyens. Dans ces conditions, le pair-expert est largement placé dans la position du géomètre-arpenteur qui, pour mesurer, doit mettre en œuvre des méthodes, des techniques et des indicateurs à la définition desquels il n'a pas participé.

2) Comme de nombreux dispositifs, elle emprunte à la gestion, à la gestion de projet et plus généralement à l'ingénierie managériale une partie de son vocabulaire (allocations de ressources, référentiel, production, qualité, indicateurs de performance, faisabilité, pilotage, tableaux de bord, échéances, livrables...), mais pas tout son vocabulaire. En effet, elle utilise volontiers les mots à la mode ainsi que le vocabulaire scientifique, pédagogique et professionnel traditionnel. On est donc davantage en présence d'un sabir (langue vernaculaire composée de plusieurs langues) que du novlangue de l'Oceania, pays à peine imaginaire d'un célèbre roman de George Orwell¹⁰, fondé sur une simplification lexicale et syntaxique de la langue destinée à favoriser la parole officielle et à empêcher l'expression de pensées hétérodoxes ou subversives (sur le vocabulaire managérial, cf. Both, 2008).

3) L'impératif gestionnaire s'incarne dans des instances, des normes, des procédures, des indicateurs (...) qui tantôt se substituent aux anciens, tantôt se complètent, tantôt se superposent.

L'AERES comme tiers spécialisé dans l'évaluation

L'idée de confier certaines évaluations à de nouvelles instances s'écartant partiellement du régime collégial, mais aussi de la formule administrative centralisée et contestée des inspections, connaît une première concrétisation à partir des années 1980, dans un contexte marqué par la réorganisation de la recherche publique (suppression de la DGRST, réforme du CNRS, succès croissant de thématiques telles la finalisation économique des recherches, la taille critique des laboratoires, le transfert des connaissances...) et par la

09. La revue *Politiques et Management Publics* contient de nombreux articles sur les différentes dimensions du NPM ainsi que sur leur mise en œuvre.

10. On fait ici référence à un débat organisé en janvier 2010 à l'EHESS dans le cadre du séminaire « Politique des sciences », au cours duquel un intervenant (Jean-Pierre Encrevê) avait qualifié ce langage de novlangue.

contractualisation des universités, mais aussi par des phénomènes d'une autre nature : poursuite de la longue marche vers l'université de masse, avec pour corollaire l'accroissement significatif du nombre d'enseignants-chercheurs (et donc « mécaniquement » du nombre de travaux, de revues, d'ouvrages, de thèses), multiplication et diversification des filières et diplômes, montée en puissance de la professionnalisation (Boure, 2007).

C'est ainsi que se crée en 1984 le Comité national d'évaluation (CNE), devenu autorité administrative indépendante par la loi d'orientation de l'éducation du 10 juillet 1989. Sa mission principale est d'évaluer l'ensemble des établissements publics à caractère scientifique, culturel et professionnel (universités, écoles et grands établissements relevant de la tutelle du ministre de l'enseignement supérieur et dans certaines conditions, établissements d'enseignement supérieur relevant d'autres ministères) et plus précisément la formation initiale et continue, la recherche scientifique et technologique et la valorisation de ses résultats, la diffusion de la culture et de l'information scientifique et technique et enfin la coopération internationale¹¹. On peut aussi évoquer le Comité national d'évaluation de la recherche (CNER) créé en 1989 sous la forme d'une structure consultative interministérielle dont la fonction première est d'évaluer la mise en œuvre et les résultats de la politique de recherche définie par le gouvernement et plus particulièrement d'évaluer périodiquement les organismes et les programmes financés par le budget civil.

Enfin, on soulignera la transformation en 2003, au sein du MESR, de la Mission scientifique universitaire (MSU) en Mission scientifique, technique et pédagogique (MSTP). Derrière le changement d'appellation, se profile un changement de braquet au niveau de l'évaluation. Divisée en 10 départements chargés de l'expertise et de la prospective, elle évalue tous les quatre ans par l'intermédiaire de ses experts (2500 environ) la politique de recherche des universités, les laboratoires, les Programmes pluri-formations (PPF), les formations et diplômes, les écoles doctorales, les candidatures à la Prime d'encadrement doctoral et de recherche (PEDR). Les critères, les synthèses des évaluations et les noms des experts sont publiés.

Selon les termes mêmes de Valérie Pécresse, « ces différents organes ont contribué à renforcer la culture de l'évaluation dans l'ensemble des universités et des organismes, mais ils restaient trop éclatés, trop spécialisés et quelquefois trop proches des formations, des institutions et des équipes qu'ils évaluaient

11. Entre 1986 et 2006, le CNE a évalué la totalité des universités françaises et plusieurs écoles dépendant de plusieurs ministères, mais aussi des sites, ce qui a permis de prendre en compte les interactions d'un ensemble d'établissements sur un territoire donné (Lyon, Normandie, Aix-Marseille, Grenoble, Montpellier, Bretagne et Pays de Loire...). En outre, il a produit des analyses concernant des activités spécifiques liées à l'enseignement supérieur (pratique du sport par les étudiants, valorisation de la recherche...) et des évaluations transversales et comparatives relatives à une discipline (géographie, sciences de l'information et de la communication...) ou à un type de formation (odontologie, 3^e cycle de médecine générale, formations juridiques de base...). Tous ses rapports (240 en novembre 2006) ont été publiés. Il semble toutefois que ses recommandations aient été assez peu suivies d'effets, tant par les tutelles que par les établissements.

pour que leur légitimité puisse être absolument incontestable. Il était donc essentiel de franchir une nouvelle étape dans la diffusion de l'évaluation en regroupant en une seule instance toutes les activités existantes et en garantissant son indépendance, sa légitimité, sa transparence, son ouverture et son efficacité » (discours du 7 décembre 2007, *op.cit.*). Et cette instance, c'est l'AERES dont la création à marche forcée, la délimitation du périmètre d'action, le mode de fonctionnement et l'économie générale du dispositif d'évaluation doivent beaucoup à Jean-Marc Monteil, alors Directeur général de l'Enseignement supérieur, puis premier éphémère Président de l'Agence (après trois mois il rejoint le Cabinet du Premier ministre).

L'AERES : une « autorité administrative indépendante » au périmètre étendu

Créée en 2006 par la loi du 18 avril de programmation de la recherche¹², l'AERES est opérationnelle depuis le 21 mars 2007, date de son installation par François Goulard, Ministre délégué à l'enseignement supérieur et à la recherche qui déclare alors dans un communiqué : « par ses évaluations, elle participera au pilotage de la recherche et de l'enseignement supérieur, en incitant tous les acteurs à s'inscrire dans un cercle vertueux d'amélioration de leur efficacité ». Comme le CNE, elle est une autorité administrative indépendante : elle fait partie de l'administration de l'État puisqu'elle agit au nom et pour le compte de l'État sans disposer de la personnalité juridique et sans être rattachée à un département ministériel. Si le terme « agence » rappelle le modèle américain, il porte aussi en lui un certain nombre d'éléments liés à la modernisation « à la française » de l'Administration et aux formes particulières que revêt en France la crise de l'État-providence.

L'AERES est totalement dédiée à l'évaluation, la décision relevant des tutelles et des établissements et organismes concernés. Elle remplace le CNE, le CNER et la MSTP et elle reprend en outre les compétences d'évaluation des unités de recherche jusqu'alors assurées par les commissions internes de chaque organisme de recherche (CoNRS, commission scientifique spécialisée de l'INSERM...). Son périmètre d'action est donc très étendu tant au niveau des établissements et organismes concernés qu'à ceux des activités de ces établissements et organismes. Concrètement, elle :

- Évalue les établissements et organismes de recherche, les établissements d'enseignement supérieur et de recherche, les établissements et les fondations de coopération scientifique et l'ANR ; les activités de recherche conduites par les unités de recherche des établissements et organismes mentionnés ci-dessus ; les formations et les diplômes des établissements d'enseignement supérieur, y compris les Écoles doctorales.
- Valide les procédures d'évaluation des personnels des établissements et organismes, et donne un avis sur les conditions dans lesquelles elles sont

12. Pour être complet, on signalera la création, également controversée, en 2005 de l'ANR dans le but explicite de refondre le système de financement public et le pilotage de la recherche.

mises en œuvre. Il s'agit bien d'une évaluation des procédures et non des personnes, celle-ci relevant toujours des sections du CNU et des commissions spécialisées des organismes de recherche.

- Participe à l'évaluation d'organismes étrangers ou internationaux de recherche et d'enseignement supérieur.

Les caractéristiques générales de l'évaluation

Elle est transversale à tous les objets d'évaluation, ce qui permet la mise en œuvre d'un dispositif général. En même temps, elle se veut conçue de façon à prendre en compte les spécificités de chaque objet évalué (établissements, unités de recherche, formations, écoles doctorales) ainsi que le contexte propre à chacun. Elle opère par vagues selon le calendrier de la contractualisation établi par académies (notion de périodicité, d'abord quadriennale, puis quinquennale).

Elle repose sur des dispositifs spécifiques construits par l'AERES « à distance des milieux professionnels concernés » (Dodier, 2009, 10) et qui s'imposent à eux. S'ils ne sont pas totalement exogènes pour ces milieux (les délégués et les experts sont formellement des pairs, certains critères sont traditionnels), ils empruntent largement au *New public management* (cf. *infra*).

Elle se veut ouverte. Cela signifie d'abord qu'elle est publique (l'AERES dit plus volontiers « *transparente* ») : la méthodologie, les procédures, les noms des experts, les rapports sont publiés, notamment sur le Web. La publicité est régulièrement mise en perspective avec l'utilité. L'évaluation serait, en premier lieu, utile pour les structures évaluées car elle permettrait de se connaître à partir de critères « objectifs », de détecter ses dysfonctionnements, d'améliorer la qualité de façon à se conformer aux standards européens, d'adapter sa stratégie et de donner des arguments aux instances pour entreprendre les réformes indispensables. Elle serait également utile aux tutelles (négociation avec les établissements et organismes, attribution de moyens...) et à la société civile (réponse aux besoins d'information des étudiants, des entreprises et des citoyens).

L'ouverture passe enfin par l'affirmation du caractère à la fois contradictoire (mais la contradiction se borne à un droit de réponse des évalués publié en même temps que le rapport d'évaluation) et évolutif, puisque l'AERES affirme prendre en compte les observations internes et externes. Cependant l'ouverture ne va pas jusqu'à la remise en question du cadre général de l'évaluation.

Elle est double et intégrée : elle repose d'abord sur une procédure jusqu'alors étrangère au monde de l'enseignement supérieur et de la recherche, mais préconisée par le *NPM* dans le cadre de « l'amélioration continue » : l'auto-évaluation des évalués, qui permet « l'évaluation interne des performances » (document en ligne). Elle se poursuit par une évaluation par des experts extérieurs relevant d'une structure collective *ad hoc* : le comité de visite pour

les établissements, les unités de recherche, les écoles doctorales et le comité de site pour les formations et diplômes (deux experts par licence ou master).

Elle est aussi intégrée, du moins en partie. En effet, si les formations et les unités de recherche sont évaluées sans interactions entre les évaluations, pour les établissements « l'AERES considère que les trois types d'évaluation ne peuvent se dérouler indépendamment les uns des autres » (document en ligne). Elle a donc mis en place une procédure spécifique : schématiquement, les évaluations des formations et des laboratoires sont transmises à la section des établissements et des rapports de site sont réalisés. Dans le même temps, des échanges d'informations sont organisés entre les trois sections. Présentée comme une évidence, cette intégration crée des interdépendances et des chaînes de responsabilités entre évalués tantôt voulues, tantôt subies : ainsi, une mauvaise évaluation d'une formation ou d'un laboratoire important peut rejaillir sur un établissement, et réciproquement.

Elle s'inscrit de plain-pied dans une démarche qualité dont l'auto-évaluation évoquée *supra* n'est qu'un élément. La Charte de l'évaluation expose « les principes qui garantissent la qualité de l'évaluation » (document en ligne). De façon plus précise, l'AERES « s'engage à mettre en œuvre un système de management de la qualité fondé sur l'approche processus (...) et à améliorer de façon continue l'efficacité de ses méthodes et procédures » (*idem*). L'Agence applique les standards européens pour l'assurance qualité dans l'espace européen de l'enseignement supérieur et elle dispose d'un *Manuel Qualité*. Enfin, elle est elle-même évaluée par une instance européenne et son rapport d'auto-évaluation est mis en ligne sur son site.

Quelques interrogations majeures

Avant même son installation, l'AERES suscite déjà de nombreuses réactions négatives dans le monde académique. De façon significative, fin 2006, le Conseil supérieur de la recherche et de la technologie (CSRT), le Conseil national de l'enseignement supérieur et de la recherche (CNESR) et la CP-CNU émettent des avis défavorables. Dès qu'elle devient opérationnelle, elle est l'objet de critiques d'autant plus virulentes que la première véritable vague d'évaluation (vague C) coïncide avec la montée en puissance du mouvement de contestation de la LRU et plus généralement des réformes en cours : multiples prises de positions publiques de personnalités, mais aussi d'enseignants-chercheurs et chercheurs anonymes via les blogs, pétitions, débats, envahissement des locaux de l'AERES, lettres collectives et individuelles, souvent mises en ligne, appelant au refus ou au boycott des comités de visite, remise en cause par plusieurs sections SHS du CNU de la pertinence de listes de revues hiérarchisées par la bibliométrie...

Le soufflet de la contestation générale retombé et les premiers bilans ayant été dressés sur telle ou telle dimension du système d'évaluation, que peut-on pointer comme interrogations majeures ? On laissera de côté des questions

dont le traitement excède l'ambition de cet article et les moyens de son auteur¹³, pour se concentrer sur quelques points qui « font problème ».

Le déni de l'élection?

L'AERES est dirigée par un organe délibérant, le Conseil, et un exécutif, le Président. Depuis la réforme constitutionnelle de 2008, ce dernier est nommé par le Président de la République, après que sa candidature ait été validée par les commissions permanentes du Parlement, compétentes en matière d'enseignement et de recherche.

L'Agence est divisée en trois sections correspondant à trois types d'évaluations (établissements, unités de recherche, formations et diplômes) dirigées chacune par un directeur. Chaque section dispose de délégués (111 en novembre 2010, adjoints compris) qui jouent un rôle déterminant dans l'organisation, la conduite et la finalisation de l'évaluation (cf. *infra*) et d'un vivier d'experts. Enfin, les délégués sont répartis en trois domaines : sciences dures, sciences du vivant et sciences humaines et sociales¹⁴.

En contradiction avec la collégialité, son organe délibérant, le Conseil, ne comporte aucun élu. Comme pour le CNE, dira-t-on. À deux différences notables près toutefois :

- Alors que dans le Conseil du CNE 19 membres sur 25 représentaient la communauté scientifique, dans celui de l'AERES 23 membres sur 25 représentent l'enseignement supérieur et la recherche, y compris privée. Toute référence à la communauté scientifique a disparu.

- Au CNE, sur les 19 représentants de la communauté scientifique, 11 membres étaient choisis sur proposition de listes de 11 noms présentées par les Présidents des sections du CNU, les Présidents des sections du CoNRS et l'Institut de France ; 3 membres étaient désignés sur une liste de 9 noms présentée par le bureau de la CPU ; 2 membres étaient choisis respectivement sur une liste de 3 noms présentée par le bureau de la Conférence des directeurs d'écoles et de formations des ingénieurs et sur une liste de trois noms proposée par les directeurs d'Institut universitaire de formation des maîtres réunis en collège (les 3 autres exerçant à titre principal des fonctions d'enseignement et de recherche dans un organisme étranger d'enseignement supérieur sont nommés après avis de l'Association européenne de l'Université). Dans l'instance délibérante de l'AERES, la proportion « d'élus » n'est plus la même.

13. Par exemple, à partir d'études de cas et dans une perspective comparatiste, on pourrait analyser les tensions qui traversent l'évaluation : AERES vs CNU ; évaluateurs vs évalués et tensions entre évalués ; quantitatif vs qualitatif ; rationalité gestionnaire vs débat démocratique ; données de management vs données déduites des pratiques d'enseignement et de recherche ; concurrence entre institutions et entre individus vs compétition intellectuelle ; excellence vs excellence républicaine (ascenseur social) et excellence intellectuelle (reconnaissance par les pairs via les productions intellectuelles)...

14. Au 30 novembre 2010, les délégués se répartissent par domaines de la façon suivante : SHS : 43 ; SD : 37 et SDV : 31. La surreprésentation des SHS est uniquement liée au nombre de disciplines relevant de ce domaine.

En effet, siègent : 9 personnalités qualifiées (mais aucun critère n'est défini et aucun avis d'une instance publique n'est requis), dont un tiers au moins issu de la recherche privée ; 7 membres ayant la qualité de chercheur, d'ingénieur ou d'enseignant-chercheur sur proposition des directeurs ou des présidents des établissements publics d'enseignement supérieur et de recherche et des organismes de recherche ; 7 membres ayant la qualité de chercheur, d'ingénieur ou d'enseignant-chercheur, sur proposition des instances d'évaluation des personnels compétentes en matière d'enseignement supérieur et de recherche (CNU et CoNRS).

Les délégués et les experts ne sont pas davantage élus, « élus d'élus » ou proposés par une instance comportant des élus, à l'exception de l'expert représentant le CNU et le CoNRS pour les seuls comités de visite des laboratoires. Les premiers sont nommés par le Président, sur proposition du directeur de section « en raison de leur compétence » (document en ligne), formule laissant entendre que l'autorité qui nomme dispose d'une grande latitude. Pour les seconds, la procédure est plus complexe, mais le résultat est le même : les critères laissent une grande liberté à l'Agence (cf. *infra*).

Qu'est-ce qui explique la place de la cooptation dans l'opacité, option qui ne semble guère en phase avec la collégialité et la transparence revendiquées ? Répondant à une question de chercheurs à propos du mode de désignation des experts, Jean-François Dhainaut, Président de l'AERES affirme : « de toute façon, cela ne faisait pas de différence que les experts soient nommés ou élus » (conférence du 8 janvier 2010). Faut-il voir dans ces propos un rejet idéologique et/ou pragmatique¹⁵ du principe de l'élection ? On y lira plus prosaïquement la croyance d'un converti au *NPM* en la faible pertinence du critère de l'élection à partir du moment où l'on est persuadé que : 1) l'Agence est (ou va rapidement devenir) un tiers extérieur réellement indépendant à la fois des évalués et des tutelles et qui finira par être accepté par le milieu professionnel ; 2) la compétence de chaque délégué et expert se mesure à ses performances attestées par son parcours et exprimées dans son CV publié sur le site, ainsi qu'à la formation reçue dans le cadre de l'AERES, tandis que son impartialité est garantie par son engagement à respecter principes et règles ; 3) la crédibilité du processus d'évaluation repose sur un dispositif à la fois conforme à la démarche qualité, objectif, performant et évolutif.

Or la légitimité et l'acceptation sociale du tiers extérieur sont encore loin d'être jouées, à tel point que d'aucuns s'interrogent de façon récurrente sur sa nécessaire réforme, quand ce n'est pas sur sa suppression. Son indépendance est contestée, non pas tant vis-à-vis des entités évaluées que vis-à-vis des hauts sommets de l'État (nomination des dirigeants, dépendance à l'égard du budget de l'État et, surtout, de la politique d'enseignement et de recherche

15. Par exemple, nommer des affidés, éviter les élus syndicaux ou relevant de listes se réclamant de mouvements tels Sauvons la recherche ou Sauvons l'université, très présents dans les sections du CNU et dans les comités des organismes de recherche.

et de pilotage de ces activités). Plusieurs années après la mise en place du processus d'évaluation, sa crédibilité est toujours sujette à controverses, y compris sur des points importants.

Et qu'en est-il de la représentativité et des compétences des délégués et des experts au regard des critères mêmes de l'AERES? Si celles des experts sont difficiles à « évaluer » (*sic!*) en raison de leur nombre et de leur diversité (l'Agence en a appelé 1 800 en 2008 et 4 500 en 2010, cf. document en ligne), celles des délégués ont pu être approchées à travers une étude réalisée à partir de leurs CV.

Bien que la mise en ligne des CV soit quasi obligatoire en raison de la publicité revendiquée par l'AERES, seuls 81 sur 111 sont accessibles à la date du 30 novembre 2010. La plupart sont construits de la même façon et comportent le même nombre de rubriques (études, carrière, principales fonctions occupées, activités d'enseignement, activités de recherche et principales publications), ce qui suggère qu'un « modèle » a été proposé et assez largement suivi dans ses grandes lignes.

Si l'on se réfère au tableau 1, les universités fournissent 68 % des délégués, contre 20 % pour les organismes de recherche (dont 12 % pour le seul CNRS). Les professeurs d'université dominent (65 %) au détriment des directeurs de recherche et surtout des maîtres de conférences (9 %) et des chargés de recherche (aucun) : la parité qui est désormais la règle dans nombre d'instances collégiales nationales et locales n'est donc pas respectée. Enfin, les délégués sont très majoritairement masculins (79 %, alors que les femmes représentent désormais 36 % des universitaires), tandis que l'Île de France est surreprésentée (38 %, alors qu'elle concentre 24 % des universitaires¹⁶)

Établissements	Statut	Zones géographiques	Sexe
Univ: 55 (68 %) Écoles ingénieurs: 5 ENS: 1 CNRS: 10 INSERM: 3 INRA: 3 CEA: 2 Collège de France: 1 Ministère culture: 1	Pr. U. : 53 (65 %) Pr. Mines: 1 MCF: 7 D.R.: 18 (22 %) Autres: 2	Île de Fr.: 31 (38 %) Régions: 50 (62 %)	Hommes: 88 Femmes: 33

Tableau 1: Répartition générale des délégués

16. Source: GESUP 2, Ministère de l'enseignement supérieur et de la recherche, novembre 2010.

Seulement 9 délégués sur 81, soit 11 %, ont exercé, hors évaluation, des responsabilités nationales, 89 % ayant des responsabilités locales relativement « lourdes » (tableau 2). Cela signifie que la plupart ont déjà connu des évaluations au titre des établissements et composantes (30 %), laboratoires (50 %) ou formations (27 %) qu'ils dirigeaient et qu'ils ont une connaissance pratique de ces institutions, de leurs atouts et de leurs problèmes. 50 d'entre eux, soit 62 %, sont ou ont été impliqués en tant que responsables ou « simples » évaluateurs dans des dispositifs d'évaluation nationaux (principalement MSTP, CoNRS et INSERM), ce qui traduit une certaine continuité avec la période pré-AERES.

Responsabilités hors évaluation*	Responsabilités dans évaluation*
Respo nationale dans org. de recherche: 9	Responsabilité nationale: 21
Président université: 4	Membre instances nationales: 29
VP univ. ou gde école, dir. ED ou UFR: 20	
Dir. laboratoire: 41	Évaluation pédagogique nationale: 3
Responsables de formations: 22	Évaluation pédagogique locale: 2

* Il s'agit de fonctions passées ou actuelles. De nombreux délégués ayant exercé plusieurs fonctions dans la même rubrique (par exemple « responsabilités dans une université »), seule la principale a été retenue. En outre, on n'a pas tenu compte des appartenances aux instances locales de recrutement (tous en sont ou en ont été membres).

Tableau 2 : Répartition par type de responsabilités

En décomposant certains critères du tableau 2 et en rajoutant d'autres indicateurs librement inspirés de Bourdieu (1984), également construits à partir des CV (cf. encadré), on peut avoir une idée, non pas de la notoriété scientifique ou sociale de chacun, mais du « capital symbolique scientifique et/ou social » de l'ensemble : il est globalement conforme à ce qui est attendu – mais sans plus – pour une instance de ce type. Si l'on établit une corrélation entre chaque indicateur et chaque CV (seuls 76 sur 81 sont exploitables), on obtient les profils suivants : « scientifiques ou plutôt scientifiques » : 34 (45 %), « sociaux ou plutôt sociaux » : 19 (25 %), « équilibrés » : 23 (30 %)¹⁷. Nous n'avons pas trouvé d'« atypiques majeurs » (marginiaux, personnes très faiblement dotées en capital scientifique ou social) : soit il n'y en a pas, soit ils sont arrivés « masqués » et tiennent sans doute à le rester.

17. Ces « résultats » doivent être observés avec prudence, non seulement parce qu'ils sont construits à partir des déclarations des délégués, que tous les CV sont pauvres en information et que certains ne sont pas remplis avec un souci d'exhaustivité, mais aussi parce qu'il existe des biais au niveau de la rédaction : par exemple, dans un milieu professionnel où la légitimité est sensée reposer d'abord sur le plan scientifique, il est possible que des délégués aient surexposé les éléments scientifiques de leur parcours.

Indicateurs de notoriété scientifique*

- **formation**: 5 ont réalisé leur thèse dans le cadre d'un laboratoire CNRS ou à l'Institut Pasteur, 18 sont passés par une Grande école (ENS: 9, Polytechnique: 1, Centrale: 1, autres: 7)
- **appartenance à une institution prestigieuse**: Institut universitaire de France (IUF): 6; Académie des sciences: 2 (membres correspondants); Collège de France: 1 (total: 9).
- **fonctions de direction dans un grand organisme de recherche**: CNRS (direction UPR ou UMR, direction – adjointe – d'un Département, présidence d'une section du Comité national): 19; INSERM: 4; CEA: 2 (total: 25).
- **direction d'une Maison des sciences de l'homme (MSH)**: 2
- **direction ou coordination de programmes de recherche nationaux ou internationaux**: 8
- **prix scientifique**: 14 (certains en ont obtenu plusieurs)
- **appartenance au jury d'un prix scientifique**: 2
- **conférences à l'étranger en tant qu'invités**: 11

* Des indicateurs comme les publications, les enseignements dans une institution prestigieuse autre que l'université ou l'organisme de rattachement, les directions de thèse et de HDR (...) n'ont pu être pris en compte en raison du caractère fragmentaire des informations.

Indicateurs de pouvoir social**

- **responsabilités internationales** (sociétés scientifiques, unions d'institutions d'enseignement supérieur et/ou de recherche, plateformes technologiques, réseaux de valorisation...): 18
- **présidence d'une section du CNU**: 3
- **fonctions importantes dans des instances nationales d'évaluation** (hors CNU): 18
- **fonctions de direction dans une université**: président: 4; vice-président CA, CS, CEVU, directeur École doctorale, directeur d'UFR: 20 (total: 24)
- **appartenance aux jurys de concours d'entrée de grandes écoles, de concours d'agrégation**: 6
- **décorations**: 5

** Des indicateurs importants (fonctions élevées dans les services centraux ou le cabinet d'un ministre, responsabilités importantes dans un syndicat ou dans une association liée à l'enseignement supérieur ou à la recherche) n'ont pu être pris en compte en raison de l'insuffisance des informations.

Il aurait été pertinent de vérifier si les délégués continuent à avoir une activité d'enseignement, de recherche et d'administration dans leur établissement ou organisme d'origine en phase avec celle qui était la leur au moment de leur désignation. On aurait ainsi pu « évaluer » si la majorité de ces officiellement « non professionnels de l'évaluation » ne sont pas devenus, le temps de leur fonction, des professionnels de fait. Mais cela aurait nécessité des outils d'investigation autres que ceux auxquels nous avons eu recours.

Pairs-experts ou ex-pairs ?

Sur le plan formel, les experts sont des pairs agissant volontairement et temporairement pour le compte de l'AERES. En effet, recrutés parmi les chercheurs et enseignants-chercheurs français et secondairement étrangers¹⁸, ils sont susceptibles d'être eux-mêmes évalués soit à titre personnel, soit du fait de leurs appartenances institutionnelles. Formés par l'AERES, ils sont en outre, tenus de respecter un certain nombre de principes et de règles (autonomie des établissements, confidentialité, respect des personnes). Leurs obligations sont mentionnées dans divers documents, tous disponibles sur le site : charte de l'évaluation, règles de déontologie de l'Agence, statut de l'expert, lettre d'engagement. Enfin, ils deviennent membres non permanents et rémunérés de l'AERES pendant la durée de la mission, ce qui peut aussi être lu à l'envers : le reste du temps, ils sont toujours (enseignants-)chercheurs.

C'est en partie sur leurs compétences que repose l'expertise « sur le terrain » et c'est à eux qu'il appartient, à partir des éléments produits par les évalués et des documents fournis par l'AERES de renseigner les grilles, puis proposer remarques, recommandations et notes. Tout au long de ce processus, l'Agence leur recommande de n'être ni censeurs, ni trop bienveillants, mais « objectifs » et « professionnels ».

Cependant, dès lors que l'on s'écarte du point de vue formel, des interrogations apparaissent.

Tout d'abord le passage de l'état d'expert potentiel appartenant au vivier à celui d'expert officiellement désigné n'est pas très clair¹⁹. Pourquoi tel pair est-il retenu pour telle évaluation et pas un autre ? Si l'on s'en tient aux dispositions officielles, la réponse va de soi : parce que le délégué l'a choisi, puisque c'est lui qui, « sous la responsabilité du directeur de section, procède au choix des experts » (*Statut du délégué scientifique*, document en ligne). En fonction de quoi ? Les critères officiels, larges et multivariés, laissent en fait une grande liberté aux délégués : compétences disciplinaires reconnues et ouverture à la multidisciplinarité, exercice de responsabilités, complémentarité des compétences avec les autres experts, fiabilité, auxquels il faut ajouter des compétences spécifiques à chaque section concernée (établissements, unités de recherche, formations et diplômes). Sans parler des critères imposés par les circonstances, comme celui de la disponibilité dont on aurait tort de penser qu'il est accessoire tant les sollicitations des (enseignants-) chercheurs sont aujourd'hui nombreuses et chronophages.

18. En 2011, l'AERES annonce disposer d'un vivier de plusieurs milliers d'experts issus de propositions d'instances diverses et de candidatures spontanées et pour lesquels l'inscription a été validée pour 4 ans et pour chaque section, par une commission d'inscription, et en dernier ressort par le Président.

19. En outre, jusqu'en 2011, l'AERES n'a jamais publié a priori pour chaque vague la liste et le CV des experts retenus. La liste d'experts était rendue publique a posteriori et les CV n'étaient pas toujours joints. Pour l'évaluation des licences et des masters de la vague C (2012), il semble que l'Agence s'oriente vers la première solution, plus conforme à l'impératif affiché de transparence.

D'ailleurs les experts ont-ils des compétences effectives dans les domaines qu'ils doivent évaluer? L'expérience montre que fréquemment des champs de recherche avérés ou des configurations pluridisciplinaires pourtant bien connues ne sont pas couverts par les comités de visite des laboratoires. Et que dire de l'évaluation des licences? Si l'AERES a mis en place depuis 2010-2011 une expertise confiée à deux experts par formation évaluée, elle demande désormais à chacun d'évaluer six mentions (contre deux auparavant), certaines pouvant relever de disciplines très éloignées de la sienne... De là à penser que pour l'AERES les contenus n'ont qu'une importance relative, il n'y a qu'un pas... que nous franchirons *infra*.

Effet prévisible du recours à un tiers extérieur, la conduite et la finalisation de l'évaluation échappent pour une part non négligeable aux experts au profit du délégué. Au cours des diverses réunions (formation, visite, séances du comité de site, réunion de synthèse...) et à travers des contacts par méls ou téléphone, le délégué utilise les différentes « casquettes » du médiateur multicartes : il rappelle régulièrement les principes et les règles (il est superviseur et parfois gendarme), explique les dimensions techniques du dispositif (il est pédagogue), légitime au besoin (il est idéologue, parfois malgré lui) et écoute les experts, y compris leurs récriminations (il est régulateur).

Son rôle ne s'arrête pas là : il se poursuit jusque dans la réalisation intellectuelle et matérielle de chaque rapport. Sur la base d'arguments techniques pour partie recevables (lisibilité pour les publics non initiés, harmonisation des styles entre les différentes évaluations, respect de l'autonomie...), il intervient systématiquement dans leur rédaction. Concrètement, il fait de nouvelles propositions « d'écriture » que le rédacteur initial hésite souvent à refuser. Ainsi amendé, chaque rapport est transmis par le délégué à son supérieur pour contrôle et éventuellement corrections²⁰.

Cependant, le principal problème vient moins de l'omniprésence des délégués que du fait que les experts sont inmanquablement conduits à mettre en œuvre un dispositif qu'ils n'ont pas construit et qui repose largement sur la mesure des performances²¹, manières de faire encore largement extérieures au monde de l'enseignement supérieur et de la recherche. Le risque est qu'ils intériorisent ce dispositif et modifient insensiblement leur rapport non seulement aux formes collégiales d'évaluation, mais encore aux pratiques

20. Pour les évaluations comportant un comité de visite, le supérieur participe à la réunion finale de synthèse, ce qui lui permet d'intervenir « sur le tas ».

21. Il serait utile de consacrer une étude spécifique aux indicateurs utilisés pour les évaluations, à leur typologie (indicateurs de caractérisation, de performances et de classements, de pilotage, de suivi des activités d'enseignement et de recherche), à leur pertinence, à leurs usages, mais aussi à leur production. Il faut en effet savoir que l'AERES ne produit pas tous les indicateurs auxquels elle a recours. Elle utilise largement non seulement ceux du Ministère, mais aussi ceux que les établissements et organismes construisent dans le cadre d'une politique de plus en plus guidée par « l'évidence gestionnaire » et son corollaire, la domination de « l'esprit gestionnaire » (Ogien, 1995), et en cela en phase avec les injonctions venant des plus hauts sommets de l'État et de l'Union européenne.

scientifiques, par exemple en prêtant plus d'attention à la gouvernance qu'à la production scientifique. Mais alors, seront-ils encore des pairs (Barthélémy, 2009) ?

« L'oubli » de l'essentiel ?

De prime abord, on pourrait penser que l'AERES est victime d'une illusion (au sens de croyance fondamentale, de prise au sérieux) sur la capacité de ce type d'évaluation à rendre pleinement compte de la richesse, de la spécificité, des projets et des difficultés de l'objet évalué. Mais est-ce véritablement une illusion, autrement dit est-ce bien cela que l'AERES cherche *in fine* à évaluer ? Sinon, comment expliquer la forte relativisation des contenus ? En effet, plusieurs experts et évalués ont déploré, ouvertement ou *mezzo voce*, l'extrême pauvreté de la discussion sur les contenus (enseignements dispensés, travaux de recherche...) et, *a contrario*, l'importance accordée à la gouvernance, aux documents comptables, à la stratégie, à la lisibilité externe (quand ce n'est pas à la communication), aux résultats souvent présentés de façon simplifiée sous la forme d'indicateurs...

Dans le même ordre d'idées, le recours préconisé à la bibliométrie pour évaluer les publications a été maintes fois dénoncé, plus particulièrement dans les SHS où la place occupée par le qualitatif et le symbolique autant que les pratiques de publication s'accordent mal avec ce type de quantification (Charles, 2008 ; Gingras, 2008). De même, le comptage des contrats « qualifiants » type ANR n'est pas susceptible à lui seul de rendre compte de la qualité du travail scientifique d'un laboratoire. Bref, pour traiter des contenus, il faut d'abord les lire (et pas seulement vérifier que les publications sont « citées » et « référencées » dans une base de données, anglo-saxonne de préférence), puis prendre le temps d'en débattre avec des spécialistes du domaine et les producteurs eux-mêmes... ce qui n'est guère possible compte tenu des conditions matérielles et de temps dans lesquelles se déroulent les évaluations. D'ailleurs, si l'AERES n'interdit pas à chaque expert de le faire, elle ne le lui demande pas.

On est donc *in fine* en présence d'une approche de ce bien public qu'est la connaissance très différente de celle qui prévalait jusqu'à ces dernières années. Cette critique est en phase avec celle du Médiateur de la République dont le dernier rapport dénonce, dans l'Administration française, le primat croissant des normes et des procédures sur les contenus, et donc sur ce qui fait l'épaisseur de la réflexion et de l'action publique.

Sur la notation

Non officiellement prévue au départ, la notation est rapidement apparue comme un élément important du dispositif, ce qui est somme toute en conformité avec la logique globale d'une ingénierie construite pour évaluer, surtout par la mesure, les performances. Elle est en outre rapidement devenue

l'élément le plus attendu par les établissements, les tutelles... et les évalués. D'abord parce qu'elle est rendue publique (elle est une des composantes du rapport, au même titre que les observations et les recommandations), ensuite parce qu'elle n'est pas sans conséquences.

Dans la pratique, les notes sont établies sur la base, d'une part, des éléments fournis par les experts, d'autre part d'une négociation, dans des conditions variables selon les types d'évaluations et les situations concrètes de chaque évaluation, entre les experts – le président du comité de visite pour les laboratoires –, le délégué et son supérieur. Elles sont toujours proposées à l'issue d'une réunion. Prenons le cas des unités de recherche. À partir de la campagne 2009-2010, l'AERES a mis en place un système permettant de donner une note globale (A+, A, B, C) et de noter, avec la même échelle, quatre niveaux : qualité scientifique et production ; rayonnement, attractivité et intégration dans l'environnement ; stratégie, gouvernance et vie de l'unité ; projet.

Premier problème, comment passer des notes de chaque niveau à la note finale ? Bien que l'AERES précise que « ni les quatre notations spécifiques, ni la notation globale ne découleront de l'application automatique d'une formule, fût-elle complexe » (document en ligne), la note globale est de fait largement conditionnée par les notes spécifiques : si un laboratoire a obtenu trois C et un B, on voit mal comment il pourrait avoir autre chose que C en note globale. Donc on attribue cette dernière en mettant en perspective des éléments quantifiés non ou difficilement comparables (comment additionner la gouvernance et la qualité scientifique ?). Cela peut s'avérer aussi délicat qu'additionner des carottes et des camions, sauf à privilégier une dimension accessoire, la couleur par exemple...

Deuxième problème, qui est source d'incompréhension de la part des évalués et de possibles conflits avec leurs établissements de rattachement, les distorsions entre le contenu et la tonalité du rapport (euphémisation des critiques) et les « mauvaises » notes. L'expérience montre que ce phénomène n'est pas rare.

Troisième problème qui commence à poindre, en tout cas officiellement, au moment où l'Agence change de Président, le mandat de Jean-François Dhainaut ayant expiré en mars 2011 : l'excessive générosité des évaluateurs. Auditionné le 4 mai 2011 par la Commission des Affaires culturelles et de l'éducation de l'Assemblée nationale, son successeur pressenti déclare « L'évaluation faite par l'Agence doit être plus contrastée. Il y a beaucoup de A+ (...) Il serait utile, par sublimation, de dégager ce qui relève du parfait, qui est sans doute plus rare, ne serait-ce que pour mettre au jour les meilleures pratiques »²². Plutôt qu'un vibrant appel à l'utilisation complète de « l'échelle des notes » de façon à ne réserver l'excellence qu'aux excellents, lesquels ne

22. Source : <http://www.educpro.fr/détail-article/h92efbff11/a/aeres> (Devant les députés, Didier Houssin défend une 'évaluation plus contrastée').

sont – évidemment ! – qu’une poignée, il faut lire dans ces propos une référence et une révérence au *benchmarking* pour lequel les « meilleures pratiques » sont des exemples dont les autres doivent s’inspirer pour s’améliorer. Cela étant, l’appréciation négative du candidat à la présidence est-elle fondée sur une réalité « empiriquement constatée » ? Pour la vague D (2010-2013), l’AERES précise que seuls 28 % des unités de recherche, toutes disciplines confondues, ont obtenu A+ (document en ligne).

Mais le problème majeur est celui des usages et mésusages des notes, et en particulier de la note globale. En effet, celle-ci peut-être utilisée par les tutelles et les établissements pour : 1) établir des classements et donc situer les entités évaluées les unes par rapport aux autres dans une logique de concurrence et de mise en visibilité ; 2) habiliter les formations et labelliser les unités de recherche. Ainsi, à l’Université Toulouse 3, tous les laboratoires notés B et C ont été entendus par le Conseil scientifique, lequel a ensuite délibéré sur leur statut. Les équipes d’accueil notées C qui entre-temps ne sont pas sabordées ou ne se sont pas intégrées à d’autres unités ont été rétrogradées « laboratoires d’université » ; 3) allouer les ressources, ce qui revient à favoriser financièrement ceux qui ont été bien notés et à pénaliser les autres ; 4) ouvrir les appels d’offres aux mieux notés : ainsi le PRES de Toulouse a considéré que seules les unités de recherche notées A+ pouvaient s’inscrire dans la dynamique de dépôt d’un projet de « Laboratoire d’excellence » (Labex) ; 5) accélérer les projets de réforme des entités en donnant des arguments à leurs instances dirigeantes : « l’évaluation doit être au service des établissements », « l’évaluation doit être un atout pour les porteurs de projet dans l’établissement » sont des propos normatifs que l’on entend régulièrement dans les réunions de l’Agence.

On ajoutera enfin que la notation est *de facto* une manière indirecte d’évaluer individuellement les enseignants-chercheurs et les chercheurs, alors même que ce n’est pas le rôle de l’AERES : en effet, que vaut le membre d’un laboratoire classé C, voire B ? Quand on lui oppose ces usages, l’AERES décline toute responsabilité en arguant, d’une part, qu’elle donne seulement un avis et qu’il appartient aux tutelles et aux établissements de prendre les décisions dans le cadre des textes en vigueur, d’autre part, que la note doit être mise en perspective avec les contenus du rapport et notamment les recommandations. Cette position rappelle étrangement celle de certains scientifiques américains qui ont mis au point la bombe atomique et qui, devant le choc occasionné par Hiroshima et Nagasaki, ont rejeté toute responsabilité sur son utilisation...

On l’aura compris, l’auteur de ces lignes se situe dans une perspective ouvertement critique. En outre, la critique porte moins sur telle ou telle dimension du dispositif que sur sa dynamique globale, sur le cadre managérial dans lequel il s’inscrit et enfin sur les discours d’accompagnement et d’autorité (et de disqualification des régimes alternatifs) dont le dispositif et le cadre font l’objet. Est-ce pour autant un plaidoyer *pro domo* pour le retour aux anciennes pratiques ? D’ailleurs, de quelles pratiques pourrait-il

s'agir ? Celles reposant sur l'inspection, quelles que soient les formes de cette dernière, ont toujours été combattues par la communauté scientifique au nom de l'autonomie. Celles du CNE n'ont guère laissé de traces auprès des évalués et des tutelles. Celles de la MSTP n'ont pas laissé de bons souvenirs, notamment en raison de leur opacité, de leur caractère non contradictoire et de la forme lapidaire des soi-disant « rapports ». Quant aux formes traditionnelles de l'évaluation collégiale, force est de constater, en tout cas dans l'Université, qu'elles concernaient davantage les enseignants-chercheurs et leurs productions que les lieux dans lesquels s'organisaient la recherche et l'enseignement (laboratoires, formations, établissements) et qu'elles étaient loin d'échapper à la critique.

Dès lors, la question de l'évaluation se pose en des termes relativement nouveaux : comment penser un régime d'évaluation général, transversal aux disciplines et aux objets évalués (établissements et organismes, unités de recherche, formations, enseignants-chercheurs), davantage indépendant et pluraliste, réellement contradictoire... et plus collégial ? Peut-on penser ce régime uniquement à partir de la critique du régime actuel, et donc en faisant l'économie non seulement des comparaisons internationales, mais aussi de la critique du mode collégial lui-même et des dérives auxquelles il a donné lieu ? Faut-il nécessairement renoncer à confier l'évaluation à une ou plusieurs autorités administratives indépendantes ? Comment penser un tel régime en l'absence « d'un contre-discours unifié émanant de la communauté universitaire ou de recherche » (Charles, 2008, 14) et dans un climat qui n'incline guère à la sérénité ? La balle est dans notre camp. Encore faut-il pouvoir et vouloir s'en saisir.



BIBLIOGRAPHIE

BALANDIER G., « Variations anthropologiques et sociologiques sur "l'évaluer" », Paris : *Cahiers Internationaux de Sociologie*, n° 128-129, 9-26, 2010.

BARTHÉLÉMY M., *Les évaluateurs de l'AERES sont-ils des pairs ? Réflexions sur le débat à l'EHESS sur le thème « L'AERES et après ? »*, 2009, <http://www.sauvonsluniversité.com/spip.php?article2313>.

BERTHELOT J.-M., MARTIN O., COLLINET C., *Savoirs et savants. Les études sur la science en France*, Paris, PUF, coll. Science, histoire et société, 2005.

BEVIR M., *Handbook of governance*, London : Sage, 2010.

BEZES P., *Le modèle de l'État stratège, genèse d'une forme organisationnelle dans l'administration française*, Paris : Sociologie du Travail, 47, pp. 431-450, 2005.

BOURDIEU P., *Homo academicus*, Paris : Éditions de Minuit, 1984.

BOURDIEU P., « La spécificité du champ scientifique et les conditions sociales du progrès de la raison », *Sociologie et Sociétés*, 7 (1), pp. 91-118, 1975,

BOURE R., *Les sciences humaines et sociales en France*, Cortil-Wodon : EME, 2007.

BOTH A., *Les managers et leurs discours. Ethnologie de la rhétorique managériale*, Bordeaux : Presses universitaires de Bordeaux, 2008.

BRUNO I., « La recherche scientifique au crible du *benchmarking*. Petite histoire d'une technologie de gouvernement », Paris : *Revue d'Histoire Moderne et Contemporaine*, n° 55 (4 bis), 2008, pp. 28-45.

COLLECTIF, « Ce qu'évaluer voudrait dire », Paris : *Cahiers Internationaux de Sociologie*, n° 128-129, 2010.

CHARLES C., « À propos de l'évaluation », *Lettre de l'ARESER*, n° 18, 2008, 14 pages.

Cités, *L'idéologie de l'évaluation. La grande imposture*, Paris : PUF, n° 37, 2009.

CRAWLEY J., « Usages de la gouvernance et de la gouvernementabilité », Paris, *Critique internationale*, n° 21, 2004, pp. 52-61.

DODIER N., *Penser un régime d'évaluation de la recherche scientifique*, EHESS, 2009. <http://www.slru.ehess.org/index.php?post/2009/03/28/un-nouveau-regime-d-objectivite>.

DUBOIS M., *Introduction à la sociologie des sciences*, Paris : PUF, coll. Premier cycle, 1999.

GINGRAS Y., *La fièvre de l'évaluation de la recherche. Du mauvais usage de faux indicateurs*, CIRST, 2008. <http://www.cirst.uqam.ca>.

LERESCHE J.-P., LARÉDO P., WEBER K., dirs., *Recherche et enseignement supérieur face à l'internationalisation*, Lausanne : Presses polytechniques et universitaires romandes, 2010.

MERTON R.K., *The sociology of science*, Chicago: University Press of Chicago, 1973.

OGIEN A., *Politiques de l'évaluation, l'évaluation comme principe politique, séminaire Les politiques de la science*, EHESS, 2009. <http://pds.hypotheses.org/284>.

OGIEN A., *L'esprit gestionnaire*, Paris : Éditions de l'EHESS, 1995.

SERVAIS P., dir., *L'évaluation de la recherche en sciences humaines et sociales. Regards de chercheurs*, Bruxelles : Academia Bruylant, 2011.

COLLECTIF, *Les universités à l'heure de la gouvernance*, Toulouse : Sciences de la Société, n° 53, 2003.

SOLLA Price (D. J. de), *Science et suprascience*, Paris : Fayard, 1972 (édition originale, 1963)

Résumé : Alors que l'évaluation comme pratique et principe est une constante de la vie académique, de nombreux enseignants-chercheurs et chercheurs accueillent avec méfiance celle qui leur est aujourd'hui imposée et qui s'incarne dans des structures nouvelles comme l'AERES. Ce qui fait l'objet des critiques les plus vives, c'est moins tel ou tel élément du dispositif d'évaluation que le cadre général en fonction duquel il a été conçu. En effet, l'évaluation collégiale a été remplacée par un régime d'évaluation largement inspiré du *New public management* qui se fonde sur l'analyse de la performance à partir de l'appréciation des résultats. C'est à une analyse critique de ce régime et de l'AERES, institution dans laquelle il s'incarne, que s'attache cet article.

Mots-clefs : Évaluation collégiale, nouveau management public, AERES, pilotage, performances, indicateurs, qualité, bilan critique

Abstract : *As the practice and principle of assessment have always been considered as a constant in university life, many researchers welcome suspiciously the newly imposed assessment system as represented by the AERES 'agency'. What is most criticized is not the system itself but the framework within which the system was developed. Indeed, collegial assessment has been replaced by a new method largely inspired from New Public Management, which analyses performance mainly in terms of results. This paper is a critical analysis of this new system and of the organization in charge of implementing it, AERES.*

Keywords : *Collegial valuation, new public management, agency, AERES, piloting, performances, indicators, quality, critical balance sheet.*

