



HAL
open science

Classer les revues de sciences humaines et sociales : comment ? pour quoi ?

Claire Lemerrier

► **To cite this version:**

Claire Lemerrier. Classer les revues de sciences humaines et sociales : comment ? pour quoi ?. La valeur de la science. Pourquoi évaluer la recherche ?, Nov 2009, Liège, Belgique. halshs-00521517

HAL Id: halshs-00521517

<https://shs.hal.science/halshs-00521517>

Submitted on 28 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Texte présenté au colloque de Liège « [La valeur de la science. Pourquoi évaluer la recherche ?](#) », à paraître (sous une forme qui sera sans doute légèrement révisée) dans la revue [MethIS](#). Les propos tenus se réfèrent à la date de rédaction du texte, le 4 février 2010.

Claire Lemerrier

Classer les revues de sciences humaines et sociales : comment ? pour quoi ?

Historienne des institutions économiques au XIX^e siècle, je ne suis pas plus légitime que quiconque pour avoir un avis sur la politique de la recherche en général, ni les classements de revues en particulier : c'est pourquoi je dois préciser tout d'abord « d'où je parle ». Il se trouve que les classements de revues m'intéressent pour deux raisons liées à mes recherches. D'une part, je pratique et promeus les méthodes quantitatives, et notamment l'analyse de réseaux, non comme une manière exclusive de faire de la recherche, mais comme un ensemble d'outils utiles. De ce fait, je n'ai pas d'opposition de principe aux opérations de mesure, pourvu qu'elles soient contrôlées et que les choix faits soient explicités ; et j'ai été confrontée à des usages, plus ou moins maîtrisés, de la mesure des citations en histoire des sciences (qui n'est au fond qu'une forme d'analyse de réseaux, le lien considéré, entre des articles, livres, personnes, revues, pays, laboratoires ou autres entités, étant le fait de citer) avant que ces usages ne deviennent partie intégrante des pratiques d'évaluation dans nos disciplines. D'autre part, ma recherche principale porte sur des formes de résolutions de conflits économiques impliquant des organisations de personnes supposées être du même métier (comme les tribunaux de commerce et conseils de prud'hommes français ou l'arbitrage proposé par nombre d'organisations professionnelles). Il s'agit d'un « jugement des pairs » sans doute par bien des points très différent de celui des universitaires, mais qui pose des questions souvent similaires, en particulier sur les liens entre expertise et proximité, ou, au contraire, extériorité ; ou encore sur les risques d'« arbitraire » et de conflits d'intérêt et l'idée de remplacer, en conséquence, le jugement direct des pairs par l'utilisation d'instruments élaborés avec ces pairs, mais qui fournissent ensuite des résultats répliquables et donc, potentiellement, ne nécessitent plus leur intervention – répertoires d'usages professionnels d'un côté, listes de revues de l'autre, par exemple¹.

Tout cela ne me donne pas une autorité particulière sur le sujet, mais explique que j'aie participé en octobre 2008, avec Emmanuelle Picard et Sylvain Piron, et en réalité sur une idée de Cécile Méadel et Pierre Mounier, à la création du carnet de recherche « Évaluation en sciences humaines et sociales » <<http://evaluation.hypotheses.org>>. L'événement déclencheur avait été la publication de listes de revues par l'AERES (Agence d'évaluation de la recherche et de l'enseignement supérieur), puis le lancement par Sylvain Piron d'un appel demandant leur retrait (<<http://appelrevues.org/>>), ce qui explique sans doute le nombre de billets consacrés aux revues et à la bibliométrie dans ce carnet, même si bien d'autres thèmes ont été abordés. C'est en partie de ce que j'ai lu à l'occasion de ma participation à ce projet que je vais faire état ici². J'y ajouterai toutefois un autre élément d'expérience : ma participation à plusieurs comités de lecture de revues, à l'animation d'une liste de

1 Il me semble de ce fait nécessaire que nous enquêtions sur la réalité du fonctionnement du jugement des pairs dans nos métiers – trop souvent évoqué en fonction de modèles idéaux ou de repoussoirs précis – et que nous le comparions à ces autres pratiques. Je ne le ferai pas ici, mais de telles enquêtes émergent aujourd'hui, tant à partir de témoignages recueillis au départ dans un but militant, au fil du mouvement universitaire français de 2009, que, par exemple, dans le récent livre de Michèle Lamont, *How professors think. Inside the curious world of academic judgment*, Cambridge/Londres, Harvard University Press, 2009.

2 On me pardonnera donc de ne pas citer en note l'intégralité des billets et surtout des textes auxquels ils renvoient, mais de conseiller à la lectrice intéressée la consultation globale du carnet, à partir des divers mots-clés pertinents ici (en particulier « bibliométrie » et « AERES », points sur lesquels je ne donnerai que peu de détails).

discussion entre collègues impliqués dans des revues de sciences humaines et sociales³, et surtout au conseil scientifique du CLEO (Centre pour l'édition électronique ouverte), précédemment nommé « comité de rédaction » de Revues.org⁴. Je ne parlerai nullement ici au nom de ce conseil et je ne trahirai aucun secret sur ses délibérations, mais j'exposerai les questions que je me suis posées et quelques convictions que je me suis forgées au fil de ces diverses expériences d'évaluation collective et de discussions sur les principes de cette évaluation.

Bibliométrie et classements : à propos de l'identification de l'adversaire et des stratégies de réponse

La scientométrie, au sens large, peut inclure toute forme de mesure de la science (budgets, brevets, effectifs... autant que publications), qu'elle soit réalisée ou non dans une visée d'évaluation, tandis que la bibliométrie peut désigner toute forme de mesure du livre : traductions ou censure, par exemple, aussi bien que nombre de publications ou de citations⁵. Quant à la mesure des citations, elle peut avoir un sens dans bien d'autres contextes que l'évaluation de la recherche – comme celui de la jurisprudence, avec la mention d'une décision d'un tribunal par un autre – et est employée depuis longtemps en histoire et sociologie des sciences, non pas pour constituer des palmarès, mais par exemple pour identifier des communautés plus précises que les disciplines, ou découpées autrement, ou encore pour réfléchir sur le processus même de citation⁶ : quel genre d'article est le plus cité, par qui et pourquoi, à quel point des auteurs sont-ils cités ensemble, la citation va-t-elle à ceux qui sont déjà les mieux dotés sur d'autres plans (effet saint Matthieu)... et même, mesure de l'effet des politiques d'évaluation, la citation va-t-elle de préférence à ceux qui ont déjà un important « facteur d'impact » ? Bref, bibliométrie ne veut pas dire scientométrie, ni comptage de citations, ni usage de GoogleScholar ou de telle ou telle base de données privée, et tout cela n'équivaut ni à une démarche d'évaluation, ni à l'établissement de classements de revues : il y a des classements qui ne sont pas du tout bibliométriques – et c'est le cas des deux plus discutés en France dans ces dernières années, celui de l'AERES et celui de l'ESF (European Science Foundation) – et des opérations de mesure de citations qui ne visent pas à distribuer ou couper des budgets, mais qui permettent de faire de la science, voire de critiquer, justement, certains abus de la (mal-)mesure dans l'évaluation. C'est ainsi au nom de ses compétences en études sur les sciences, mais aussi plus précisément en

3 La liste revues_shs@cru.fr, dont la création est liée à une série de réunions entre responsables de revues en 2005, avec, déjà, pour élément déclencheur la publication d'un classement, alors réalisé par le CNRS. Voir Claire Lemerrier, « Les revues de sciences humaines et sociales prennent la parole », *La revue des revues*, n° 38, 2006, p. 95-98 et Emmanuelle Picard, « Le CNRS et les revues-2004 (opus 1) », 2008, <<http://evaluation.hypotheses.org/39>>. Toutes les url données dans cet article correspondent à une consultation du 3 février 2010.

4 Sa composition est décrite sur <<http://cleo.cnrs.fr/103>> et son fonctionnement est présenté par Delphine Cavallo, « Revues.org : l'invention de l'édition électronique scientifique, entre libre accès et modèle économique pérenne », *Mémoires du livre/Studies in Book Culture*, volume 1, n° 1, 2009, <<http://id.erudit.org/iderudit/038638ar>>.

5 Voir par exemple Michel Callon, Jean-Pierre Courtial et Hervé Penant, *La scientométrie*, Paris, PUF, 1995 ou la collection de la revue *Scientometrics* <<http://www.springerlink.com/content/101080/>>; sur la bibliométrie, Alain Vaillant (éd.), *Mesure(s) du livre*, Paris, Bibliothèque nationale, 1992.

6 Entre mille exemples, je renvoie très subjectivement à Gregoy A. Caldeira, "Legal Precedent: Structures of Communication between State Supreme Courts", *Social Networks*, vol. 10, n° 1, mars 1988, p. 29-55 ; Douglas R. White and H. Gilman McCann, "Cites and Fights: Material Entailment Analysis of the Eighteenth-Century Chemical Revolution", in Barry Wellman and Steven Berkowitz, *Social Structures: A Network Approach*, Cambridge, Cambridge University Press, 1988, p. 380-399 ; Matthew L. Wallace, Yves Gingras & Russell Duhon, "A new approach for detecting scientific specialties from raw cocitation networks", 2008, <<http://arxiv.org/abs/0807.4903v1>>, Vincent Larivière, Benoît Macaluso, Éric Archambault & Yves Gingras, "Which Scientific Elite? On the Concentration of Funding, Productivity and Citations", 2009, <<http://www.ost.uqam.ca/Portals/0/docs/articles/2009/ISSI%20Paper%20Concentration%20Final.pdf>>, Vincent Larivière & Yves Gingras, "The impact factor's Matthew effect: a natural experiment in bibliometrics", 2009, <<http://arxiv.org/abs/0908.3177>>.

scientométrie, qu'Yves Gingras est régulièrement intervenu dans les débats sur l'évaluation pour critiquer tant la construction que les usages de différents classements.

Critiquer les instruments de classement existants, ou critiquer les classements ?

Il me semble qu'il est crucial de prendre le temps de faire ces distinctions, de s'intéresser à la façon dont sont réellement constitués les classements et d'autres indicateurs, si l'on décide par ailleurs de s'y opposer, soit justement du fait de ces modes de construction, soit plutôt des usages qui en sont faits ou des effets pervers qu'ils peuvent entraîner. Le refus global de la mesure (sous prétexte que certaines choses ne seraient pas mesurables), ainsi que certains autres arguments déjà employés dans le débat, me semblent être trop aisés à démonter, ou bien décalés par rapport à ce que je considère comme le fond du problème.

Il en va ainsi de l'idée que les sciences humaines et sociales occuperaient une position particulière, que c'est peut-être leur spécificité qui devrait amener à ne pas leur appliquer la bibliométrie ou les classements. Outre qu'il faudrait discuter des lieux exacts où passent les frontières – rarement alignées de façon simple sur les divisions humaine et sociale *vs.* dure, ou fondamentale *vs.* appliquée –, l'argument méconnaît le travail de critique des indicateurs précisément mené par ceux qui y sont confrontés depuis plus longtemps que nous, y compris depuis des positions institutionnellement ou économiquement fortes : ainsi, pour ne prendre que deux exemples, de plusieurs associations internationales de mathématiques, ou encore de chercheurs en médecine discutant de l'impact de la bibliométrie sur une recherche très appliquée et légitime⁷.

Il est vrai que certaines caractéristiques communes à au moins une partie des sciences humaines et sociales (mais pas toujours limitées à elles) ont pu rendre plus absurdes qu'ailleurs l'application irréfléchie, par exemple, de facteurs d'impact calculés à partir des *Social Sciences* ou *Arts & Humanities Citation Index(es)*. Il en va ainsi du rôle des livres, des publications dans d'autres langues que l'anglais, d'une relative lenteur à citer (c'est-à-dire à la fois de citations relativement rares de travaux très récents et relativement fréquentes de travaux anciens) ou encore du renvoi par certaines citations à des sources plutôt qu'à de la bibliographie. Il est très utile de prendre en compte ces dimensions si l'on veut comprendre ce que mesure exactement un facteur d'impact, par exemple – notamment pour déconseiller son usage dans une situation particulière.

Le problème est toutefois que ce type de critique peut entraîner, et de fait entraîne déjà, une réponse sous forme d'affinements méthodologiques ou de construction de nouvelles sources – de la même façon que des critiques moins spécifiques des sciences humaines et sociales, évoquant la prise en compte des auto-citations dans les indices, l'impact du nombre de co-signataires d'un article, le poids exagéré appliqué aux citations d'un seul article pour juger de la production globale d'une personne, ou encore la faible couverture des bases de données ISI ont déjà donné naissance, ou pourraient donner naissance, à de nouveaux indicateurs, comme l'h-index, à de nouvelles bases de données, ou à des utilisations d'autres bases existantes (comme celle indexée par GoogleScholar). De façon encore plus triviale, nombre d'indicateurs bibliométriques fondés sur d'immenses bases et un calcul automatisé posent des problèmes de reconnaissance des individus (homonymie entre chercheurs, désignations différentes d'une même université), voire de typographie (accents, cédilles...) qui peuvent nettement biaiser les résultats, en particulier à un niveau fin (niveau d'analyse que les concepteurs des indicateurs eux-mêmes déconseillent en

7 Joint Committee on Quantitative Assessment of Research Citation Statistics, *A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS)*, 2008, <<http://www.mathunion.org/fileadmin/IMU/Report/CitationStatistics.pdf>> ; Julien Mancini, Stefan Darmoni, Hervé Chaudet & Marius Fieschi, « Le paradoxe de la « T2A bibliométrique » SIGAPS : un risque d'effet délétère sur la recherche française ? », *La Presse Médicale*, volume 38, n° 2, 2009, p. 174-176 <<http://www.em-consulte.com/article/199681>>.

général, entre autres pour cette raison, d'explorer, mais qui est celui visé par les politiques d'évaluation actuelles...). Une des réponses à ce problème, en France, depuis quelques années, a été la multiplication à la fois des groupes de travail et des instructions visant à standardiser les mentions d'institutions dans les signatures d'articles, et des regroupements institutionnels visant entre autres à atteindre une taille critique bibliométrique.

Plus généralement, du côté des sciences humaines et sociales, de l'Europe et/ou de la France, se sont multipliées les tentatives de construction de corpus, de classements ou d'indicateurs donnés comme concurrents ou complémentaires des plus utilisés – de la même façon que chacun crée son classement des universités pour tenter de détrôner celui de Shanghai. Les classements successifs qui ont fait réagir nombre de collègues en France, ceux du CNRS (Centre national de la recherche scientifique) en 2004, et plus récemment de l'ESF et de l'AERES, se donnaient d'ailleurs explicitement de tels buts, que l'on retrouve encore tout récemment, par exemple, dans le manifestement très lourd projet *JournalBase*⁸.

Utiliser autrement les instruments ?

Le point commun de ces critiques et de ces réactions me semble être l'acceptation de l'idée qu'il existerait une hiérarchie unique de la valeur scientifique (à la fois entre individus et entre revues), éventuellement utilisable, au moins comme un critère parmi d'autres, comme base de prise de décision, mais que les outils existants mesureraient mal cette valeur. Il n'est pas difficile de glisser de cette idée à la volonté d'améliorer les outils, ou d'en construire d'autres, pour mettre enfin au premier rang ceux qui le méritent réellement... surtout si la seule position adverse présente dans le débat est celle qui nierait, au fond, toute notion de différences de qualité dans la science, position toujours précaire pour une activité à la fois financée sur fonds publics et attentive à fixer ses propres frontières vis-à-vis d'éventuels charlatans ou simples vulgarisateurs. Nous sommes sans doute tous d'accord sur le fait qu'en sciences (même humaines et sociales), tout ne se vaut pas, et nous acceptons en permanence ce postulat pour évaluer non seulement les étudiants, mais les projets de thèse, les candidatures à des postes, les articles, etc. Nous serions sans doute également à peu près tous d'accord pour refuser une politique scientifique fondée principalement sur le classement de Shanghai ou l'*impact factor* (à l'échelle des disciplines, des revues ou des individus). Mais cela laisse en réalité ouverte toute une gamme de positions et d'arguments possibles et non équivalents.

Ainsi, des pairs peuvent se saisir des indicateurs existants, dans leur pluralité, pour souligner que la « grandeur » des scientifiques n'est pas unidimensionnelle, chacun(e), suivant sa personnalité, sa formation, son insertion institutionnelle, sa spécialité, mais aussi les phases de sa carrière, investissant plus ou moins chaque facette du métier (que l'on peut découper à l'envi entre différentes formes d'enseignement, de vulgarisation, de production d'outils pour la recherche, de recherche « de terrain », de théorisation, de publication, de travail organisationnel, de participation au jugement des pairs...), avec plus ou moins de rapidité, de loquacité, de travail individuel ou collectif... cette diversité de styles se retrouvant bien sûr si l'on se concentre sur les seules publications. Bref, si un chiffre ne dit pas grand chose, un faisceau de chiffres peut donner une première idée d'un profil de chercheur/se, aider à objectiver des intuitions venues à la lecture d'un CV, fournir au fond une base de discussion collective entre pairs qui, sans être dupes des classements, en font un élément de décision parmi d'autres. Ainsi, taper « Claire Lemerrier » dans *Publish or Perish* ne dit à peu près rien de ma recherche ni même de mon statut ou de ma réputation, si l'on en reste au h-index (sinon qu'il permet de faire des hypothèses assez justes sur la durée de ma carrière) ; en revanche, les autres informations fournies, par exemple le fait que ce score tient en grande partie à ma collaboration avec un sociologue enclin à publier à la fois en anglais et en français, au sein de

8 Michèle Dassa, Christine Kosmopoulos et Denise Pumain, « JournalBase. Comparer les bases de données scientifiques internationales en sciences humaines et sociales (SHS) », *Cybergeo : European Journal of Geography*, document 484, mis en ligne le 8 janvier 2010 <<http://www.cybergeo.eu/index22862.htm>>.

collectifs à géométrie variables et sans se refuser à l'auto-citation, peuvent être pertinentes pour comprendre ma position dans le champ académique français.

L'idée de collectifs de pairs se saisissant en connaissance de cause de telles informations n'a rien d'utopique, puisqu'elle est promue depuis plusieurs années par Yves Langevin en référence à son expérience au sein de la section « sciences de la planète et de l'Univers » du Comité national du CNRS (chargé notamment des recrutements et des promotions des chercheurs CNRS dans cette spécialité)⁹. Cependant, il a souvent été rappelé que, du point de vue des non-chercheurs, voire des chercheurs, chargés de la prise de décision, *any number beats no number* (au point que je suis incapable d'identifier la source de l'expression, qui semble avoir été introduite dans le débat qui nous occupe par Yves Gingras) ; on pourrait ajouter qu'un chiffre unique bat inmanquablement une série de chiffres, ou plutôt que la tendance sera de les additionner pour aboutir à un classement univoque¹⁰, plutôt que d'observer avec intérêt que nul n'étant le premier partout, un choix reste à faire sur les éléments de profil que l'on privilégie.

À l'attrait général des chiffres, notamment parce qu'ils peuvent ôter une partie du poids de la responsabilité d'une décision à celui qui affirme s'appuyer sur de tels indicateurs considérés comme objectifs, s'ajoute en outre souvent la volonté d'éviter, précisément, le jugement collectif et délibératif des pairs, pour des raisons explorées dans les autres exposés de cette journée. Celles-ci mêlent sans doute à des degrés divers la croyance sincère dans la résolution des problèmes d'arbitraire ou de conflit d'intérêt par l'usage d'un instrument produisant des résultats répliquables, la méfiance de principe envers les scientifiques et/ou les élus ou les syndicalistes, et le constat du temps pris et du coût occasionné, à l'échelle individuelle ou organisationnelle, par un travail sérieux de jugement collectif. À cet égard, la participation à des jurys de recrutement de maîtres de conférences en France m'a amplement convaincue que l'usage systématique d'indicateurs qui, pour ne pas être directement bibliométriques, n'en étaient pas moins largement hors de propos, comme le fait d'être agrégé-e de l'enseignement secondaire ou d'avoir publié au moins un article dans une revue bien connue était moins lié à une croyance dans l'intérêt de ces qualités pour un-e maître-esse de conférences qu'au manque de temps donné, ou pris, pour examiner les dossiers : de tels critères faciliteraient tout simplement un premier tri. Il me semble toutefois que ce manque de temps, sur lequel nous achoppons tous, ne doit pas être un argument pour s'en remettre à un certain type de classements de revue, à la fois parce qu'ils posent des problèmes fondamentaux et parce que leur fabrication est elle-même chronophage.

Des effets de l'observation sur l'observé

Un dernier point me semble trop souvent absent des discussions sur les citations et autres formes de scientométrie : le fait que les scientifiques sont tout à fait capables de comprendre la construction et les effets de ces instruments, et donc d'adapter leurs comportements en conséquence – à la fois, notamment, comme auteurs d'articles et comme membres de comité de lecture¹¹. À l'extrême, si l'on

9 On peut par exemple écouter Yves Langevin, « Évaluation par les pairs, bibliométrie et numérologie vues par un président de section Sciences de la Planète et de l'Univers », Colloque « Évaluer, dévaluer, réévaluer l'évaluation », École normale supérieure, 30 novembre 2007, <<http://www.diffusion.ens.fr/index.php?idconf=1878&res=conf>>.

10 « Jean-Louis Briquet s'est interrogé sur l'idée de multiplication des indicateurs en rapportant comment un mouvement avait été engagé au sein de l'AERES pour promouvoir l'usage de quatre notes, pour empêcher que les décideurs n'utilisent que la seule note globale attribuée par les comités de restitution. Malheureusement, il a indiqué que cette tentative s'était soldée par la création d'une cinquième note dite « globale » synthétisant les quatre autres, illustrant par là la difficulté d'introduire de la complexité dans des institutions qui cherchent à simplifier les processus de décision. » Daniel Benamouzig et Giovanni Prête, « Évaluer la sociologie : Vers de nouvelles normes d'évaluation scientifique ? » (compte rendu d'une journée d'étude), in Dossier de l'ANCSMP sur les classements de l'AERES, n° 2, décembre 2008 <http://ancmsp.apinc.org/IMG/pdf/Dossier_AERES_2-ANCSMP.pdf>

11 Si le point me semble avoir été peu présent dans les débats auxquels j'ai assisté, il a été exploré par certains chercheurs. Voir notamment Sylvain Piron, « Lisons Peter Lawrence, ou les implications morales de l'évaluation

tirait au sort une liste de revues « excellentes », elles le deviendraient sans doute au bout d'un moment, parce que tout le monde essaierait d'y publier¹²... On pourra voir dans cette adaptation le but ultime de la manœuvre de mesure, visant simplement à inciter à de nouveaux comportements, en suivant l'analyse par Philippe Büttgen des discours de Nicolas Sarkozy proposée lors de la présente journée – en reconnaissant toutefois que ce but n'est pas consciemment poursuivi par nombre de ceux qui mettent ces politiques en œuvre. On pourra aussi y voir une manière individuelle ou collective de survivre sous le nouveau régime – comme on peut constater qu'il y est généralement possible de défendre rhétoriquement le lien entre un projet de recherche très fondamental et une application à la mode, s'il en va d'un budget. Nicolas Dodier compare cette « dissociation » à celle qu'expérimentent les ouvriers soumis au taylorisme lorsqu'ils parviennent à coller *a minima* aux procédures prescrites, tout en construisant par ailleurs un autre sens du travail bien fait, au sein de collectifs plus ou moins clandestins¹³.

Enfin, on pourra évoquer les effets concrets qu'une évaluation fondée sur les publications et/ou les citations peut avoir sur les pratiques. « Soyons francs : nous savons tous faire cela, découper un article en deux, trois, quatre articles et le faire traduire en deux, trois, quatre langues pour ajouter autant d'items à nos listes de publications. Productivité en hausse, créativité zéro. Notre schizophrénie de chercheurs est là : nous renâçons – également sous la forme de la contestation infinie des « critères » – mais nous savons faire, et nous savons que ce n'est pas cela qu'il faut faire. »¹⁴ On peut penser aussi à la mise en place des groupes de chercheurs se citant mutuellement ; à la non citation délibérée des travaux de chercheurs concurrents ; à la pression exercée par certaines revues pour que l'on cite des articles publiés précédemment par elles ; et plus globalement au fait, pour les revues, d'assumer le rôle qui leur est assigné, celui de valider la qualité des articles, et indirectement d'accélérer ou freiner les carrières, plutôt que de donner à lire ces articles à un public. Certes, tout cela existait déjà en bonne partie ; je ne prétends pas que l'usage de la citation n'était guidé que par la volonté d'indiquer la source de données réutilisées ou d'inspirations théoriques, ni que les revues choisissaient les textes en fonction d'une ligne éditoriale et en essayant d'améliorer les articles, sur le fond et la forme, avec les auteurs, sans aucune arrière-pensée quant à leur rôle de *gatekeepers*. Mais il me semble constater des préoccupations plus conscientes et plus suivies d'effets quant à la longueur et à la composition des CV chez mes collègues, et, du côté des revues, une réticence à participer à l'amélioration des textes (comme relecteur ou éditeur) en vertu du fait que l'auteur serait le seul à « profiter » de la qualité de l'article obtenu *in fine*. Mon expérience m'incite ainsi à penser que les effets des instruments de mesure sur les comportements posent de réels problèmes non seulement pour le bien-être des chercheurs mais surtout pour l'avancement de la science – dans la mesure où je souhaite que la cumulativité, la mention des références, le fait de rendre publics les résultats, de le faire de façon claire, cohérente et vérifiée continuent à faire partie des idéaux régulateurs de notre métier. La capacité des chercheurs à s'adapter aux formes de l'évaluation devrait être considérée non comme une consolation, mais bien comme une des principales raisons de s'opposer à l'usage mécanique, voire à l'usage tout court, de tels instruments.

Listes et classements de revues : à quoi bon ?

On aura compris que le problème est moins, pour moi – comme pour bien d'autres –, dans

bibliométrique », décembre 2008, <<http://evaluation.hypotheses.org/229>> et Peter Lawrence, "The Politics of Publication", *Nature*, 422, 2003, p. 259-261 <<http://www.mrc-lmb.cam.ac.uk/PAL/pdf/politics.pdf>>.

12 Sophie Roux, « Comment je me suis disputée avec les classificateurs », décembre 2008, <<http://evaluation.hypotheses.org/283>>.

13 Nicolas Dodier, « Penser un régime d'évaluation de la recherche scientifique », mai 2009, <<http://evaluation.hypotheses.org/452>>.

14 Philippe Büttgen et Barbara Cassin, « Les dégâts de l'évaluation. L'État schizophrène et l'évaluation de la recherche », *La Vie de la recherche scientifique*, 378, 2009, p. 30-31. <<http://www.snscs.fr/IMG/pdf/Vrs378.pdf>>

l'instrument, et notamment dans la façon exacte dont il est construit, que dans l'usage qui en est fait, en particulier dans la réduction de la « grandeur » scientifique à une échelle unique, et dans le choix de ceux qui en font usage, en particulier dans le sort fait au jugement délibératif des pairs. Je voudrais préciser cette proposition bien abstraite (et qui peut de ce fait sembler consensuelle) en revenant sur quelques épisodes, et quelques expériences personnelles, concernant le classement, ou l'établissement de listes, de revues. Des cas que je vais évoquer, la bibliométrie sous toutes ses formes est à peu près absente ; en outre, les listes qui firent réagir nombre de collègues avaient été établies par d'autres collègues. Cela souligne que le problème est plus complexe qu'une opposition entre bibliométrie menée par des non-spécialistes dans un but de pilotage et jugement des pairs. Je ne prétends pas en proposer une analyse définitive, mais plutôt poursuivre une réflexion personnelle : comment suis-je passée de l'idée générale qu'il serait bien de reconnaître que certaines revues fonctionnent plus sérieusement que d'autres à l'opposition à la publication de toute forme de liste de revues considérées comme de qualité¹⁵ par l'AERES ?¹⁶ Comment puis-je concilier cette position avec la pratique régulière, au sein du Conseil scientifique du CLEO, d'un tri entre des objets éditoriaux acceptés comme revues scientifiques de sciences humaines et des sociales et d'autres qui ne le sont pas ?¹⁷

Pourquoi des listes de revues ?

Les listes de l'AERES ont déclenché un tollé en France, en particulier sur la base de comparaisons avec d'autres listes, ou avec un certain sens commun disciplinaire, qui tendaient à montrer que certaines revues étaient absurdement classées (trop haut ou trop bas)¹⁸ : pour beaucoup, c'était le fait que les listes ne soient pas les bonnes, plutôt que leur existence même, qui posait un problème. Le débat s'est en outre focalisé sur le fait que ces listes proposaient des classements, en général en A, B, C, en pratique traduits en « A et B » vs. « C et non classé » pour l'évaluation des laboratoires. Des réactions antérieures aux pratiques de la section d'économie et gestion du CNRS, qui publiait déjà depuis plusieurs années un classement à base de nombre d'étoiles (constitué par la section, sans base bibliométrique) sur lequel commençaient à être fondées des primes, en particulier pour des enseignants en écoles de commerce¹⁹, ont été réinvesties dans ce débat. Tout cela a conduit nombre de participants à défendre le principe d'une liste unique, c'est-à-dire d'un classement binaire (revues classées vs. non classées), souvent nommée « périmètre de scientificité ». La violence symbolique inacceptable, ou la prétention d'objectivité absurde, aurait donc résidé dans le fait d'établir une hiérarchie interne, tandis que le fait de définir ce qui est une revue scientifique de

15 En réalité, une des spécificités des listes de l'AERES est qu'il n'est à aucun moment clairement spécifié ce que sont supposées être les revues listées (lorsque cette précision est donnée, ce n'est que discipline par discipline). Je m'aventure donc à parler de qualité en espérant choisir un terme assez neutre. Ce que les listes sont supposées dire est en réalité défini, de façon circulaire, par l'usage qui en est fait : publier dans une revue classée A ou B sur ces listes permet d'être considéré comme « publiant », dans le cadre d'une évaluation du laboratoire (et *de facto* de l'individu), tandis que publier ailleurs ne le permet pas.

16 On trouvera une réflexion argumentée de façon bien plus serrée sur un – supposé – paradoxe très similaire sous la plume de Sophie Roux, art. cit.

17 J'ai déjà exposé cette interrogation dans un billet qui a donné lieu à des commentaires intéressants et dont je reprends ici les arguments, n'ajoutant que quelques références et éléments de contexte : Claire Lemerrier, « À propos d'un « périmètre de scientificité » pour les revues de SHS », mars 2009, <<http://evaluation.hypotheses.org/396>>. Ce texte avait au départ été écrit dans la perspective d'une réunion du comité éditorial du *Mouvement social*, dont plusieurs animateurs étaient très impliqués dans la défense d'un « périmètre de scientificité » contre les listes de l'AERES : voir « Déclaration du Comité éditorial de la revue *Le Mouvement Social* », mise en ligne le 15 janvier 2009, <<http://mouvement-social.univ-paris1.fr/document.php?id=1410>>.

18 Voir notamment bon nombre des articles référencés par la page « évaluation » de l'Association française de science politique, et qui ne portent pas seulement sur cette discipline <<http://www.afsp.msh-paris.fr/observatoire/etudes/evaluation/evaluation.html>>.

19 Florence Audier, « L'évaluation et les listes de revues », *La vie des idées*, septembre 2009, <<http://www.laviedesidees.fr/L-evaluation-et-les-listes-de.html>>.

sciences humaines et sociales (SHS) et ce qui n'en est pas une aurait plutôt des vertus, par exemple, selon le géographe Éric Verdeil, celle de « constituer un repère susceptible de structurer le champ des SHS ou du moins certains de ses sous-champs où les pratiques de publication se revendiquant de la recherche sont très confuses », « d'identifier clairement des lieux où la publication procède d'une validation collective et d'un débat ». Le corollaire serait alors « de défendre une conception large de la dite liste et de revendiquer que sa constitution relève de critères transparents [...] afin d'éviter que leur établissement soit pollué par des considérations de réputation, de hiérarchies implicites. »²⁰ L'historien Gabriel Galvez-Behar plaide pour le même type de listes²¹ en souhaitant qu'elles soient constituées par de grandes associations disciplinaires : un modèle qui n'existe pas pour l'histoire en France, mais qui serait par exemple celui de l'Association française de sociologie et surtout des organisations nord-américaines, beaucoup plus directement régulatrices, calquées sur le modèle des ordres professionnels. Constituer une liste de revues ne serait alors qu'une partie du travail de délimitation mené par la profession.

Ces réactions, à mon sens, pointent un mouvement, plus général et en bonne partie générationnel, en faveur d'une professionnalisation des sciences humaines et sociales, au sens où elles se détacheraient à la fois de l'image d'une pratique d'amateurs (se traduisant par le fait de s'accommoder de thèses non financées, de l'absence de bureaux, etc.), ou trop peu distincte de la vulgarisation (un point sensible par exemple dans l'édition de livres²²) et d'un monde régulé uniquement par des relations interpersonnelles (l'enjeu de l'anonymisation des articles pour évaluation étant de ce fait, par exemple, essentiel, comme, à une autre échelle, la lutte contre le « localisme » dans les recrutements universitaires²³).

Or cette exigence de professionnalisation est confrontée à une explosion du nombre d'objets éditoriaux pouvant aspirer à se faire reconnaître comme revues de sciences humaines et sociales, explosion décrite par plusieurs rapports (et dans l'intervention de Gabriel Galvez-Behar) et ayant sans doute à voir tant avec la croissance du nombre de chercheurs qu'avec la volonté de centres de recherche de se valoriser en ayant leur propre publication, fût-elle non diffusée, voire avec la facilité supposée de créer une revue électronique plutôt que sur papier. Quoi qu'il en soit, cette abondance, à laquelle j'ai été confrontée avec une certaine surprise dans mes premières années au comité de rédaction de Revues.org, peut en effet donner envie, ou rendre nécessaire, de constituer un outil d'orientation – voire, et c'est le risque de la démarche, de réduire la diversité au nom de la lisibilité²⁴. C'est pourquoi des labels existent déjà, procédant de façons diverses à une sélection entre revues incluses ou exclues ; c'est pourquoi aussi certains défendent l'établissement de listes générales par disciplines, en évoquant tout particulièrement deux situations de doute fréquentes : celle de l'auteur, et particulièrement du jeune doctorant ou docteur, se demandant où il est pertinent de publier (une liste devant alors permettre d'égaliser les chances, si la carrière ultérieure est en partie déterminée par cette publication, entre ceux qui sont déjà bien intégrés dans la communauté et les autres) ; celle du recruteur perplexe devant un CV qui évoque de nombreuses publications dans des revues inconnues de lui, et se demandant comment en juger.

20 J'extrai ces citations d'un commentaire de mon billet précité de mars 2009, parce que celui-ci me semble particulièrement construit et intéressant et non pour faire de son auteur l'emblème d'une position différente de la mienne.

21 Gabriel Galvez-Behar, « Faut-il classer les revues en sciences humaines et sociales ? Dix années de controverses françaises (1999-2009) », « L'évaluation de la recherche en sciences humaines et sociales. Approches politiques, historiques et méthodologiques », 16 novembre 2009, Université de Louvain : une présentation résumée est en ligne <<http://ggb.ouvaton.org/spip.php?article33>> et j'ai pu discuter de ses conclusions avec l'auteur en janvier 2010.

22 Bruno Auerbach, « Production universitaire et sanctions éditoriales. Les sciences sociales, l'édition et l'évaluation », *Sociétés contemporaines*, 74-2, 2009, p. 121-145.

23 Voir le dossier « Le localisme universitaire » de *La vie des idées* (avril-mai 2008), <<http://www.laviedesidees.fr/+Le-localisme-universitaire+.html>>.

24 Sylvain Piron, « "Y a-t-il un Nature ou un Science en SHS" », octobre 2008, <<http://evaluation.hypotheses.org/128>>.

Listes, modes d'élaboration et finalités

En France, différents dispositifs de sélection de revues existent déjà depuis plus, et parfois nettement plus, de cinq ans – on pourrait d'ailleurs s'étonner qu'ils ne soient pas plus mobilisés pour appuyer un jugement dans les deux situations précitées. Deux d'entre eux sont essentiellement liés au financement des revues. Le CNRS soutient un certain nombre de revues par l'attribution de postes ou demi-postes de secrétaires de rédaction, d'édition, de cartographes, etc. et/ou de subventions ; ces revues peuvent se réclamer du soutien de l'institution, depuis quelques années formulé en termes de « labellisation »²⁵. La liste actuelle comprend 177 titres pour l'ensemble des sciences humaines et sociales – donc extrêmement peu par rapport aux listes de l'AERES et même à la couverture des portails de revues. Label et soutien financier sont accordés sur la base d'une demande des revues, qui doivent déposer un dossier transmis aux sections (par discipline) du Comité national : les mêmes organes, composés à la fois de collègues élus et nommés, qui sont chargés notamment du recrutement et des promotions. Le site du CNRS souligne que l'obtention d'un tel soutien peut aider au « rayonnement » de la revue et de ceux qui y publient et permettre de décrocher d'autres financements ; il affirme aussi que les critères de labellisation croisent « qualité scientifique et reconnaissance par la communauté, transparence et rigueur de gestion, renouvellement et dynamisme des instances éditoriales ». Si l'on peut en effet imaginer que les membres du Comité national utilisent leur connaissance du contenu et de la réputation des revues acquises dans leur pratique scientifique (ou ont recours à leur bibliothèque ou au web pour s'informer avant d'évaluer), il faut souligner que le dossier rempli par les revues demande de plus en plus des informations standardisées, toutefois sur des éléments assez variés (budget, tirage, part d'abonnés et d'auteurs étrangers, d'articles refusés, statistiques de fréquentation en ligne²⁶...) pour qu'on puisse espérer la mise en valeur de revues qui sont « grandes » de manières différentes.

Le Centre national des lettres subventionne également un certain nombre de revues de sciences humaines et sociales, par des aides ciblées (pour la traduction ou la numérisation) ou plus générales ; ses critères d'attribution, qui ont été réformés et critiqués dans les dernières années, tendent à favoriser les revues qui sont lues hors du monde des sciences humaines et sociales, et accordent une grande place à la situation financière (dont les responsables de revues imaginent qu'elle ne doit être ni trop bonne, ni trop mauvaise pour qu'une aide soit accordée – je n'ai pas d'information sur la crédibilité de ces anticipations). Une tension peut ainsi apparaître entre des critères internes au monde de la recherche et ceux d'une institution qui subventionne aussi (et surtout) des revues de création littéraire et n'envisage pas forcément de la même façon le rôle du support « revue »²⁷. La commission (unique) chargée des décisions en matière de sciences humaines et sociales n'en est pas moins quasi exclusivement composée d'universitaires.

À ces deux instances composées de pairs et mobilisant des critères explicites un peu différents

25 Cette politique de soutien pour 2009 est présentée sur <http://www.cnrs.fr/inshs/recherche/docs/SHS-soutienrevues2009.pdf> et la liste de revues correspondante est donnée sur <http://www.cnrs.fr/inshs/recherche/docs/soutien-revues2009.pdf>.

26 Pour revenir sur l'adaptation des acteurs à la mesure, on peut souligner que ces nombres sont assez aisément manipulables, dans la mesure où ils mobilisent des catégories beaucoup moins évidentes qu'elles ne le semblent de prime abord. Il en va ainsi non seulement du budget (coûts mutualisés entre revues par l'éditeur, intégration ou non des heures de travail « bénévoles »...), mais aussi par exemple de la notion d'auteur étranger (suivant sa langue d'expression, son appartenance institutionnelle, voire la consonance de son nom ?) ou encore d'article refusé (faut-il compter les propositions quelque peu fantaisistes – et dans ce cas, pourquoi ne pas en susciter pour augmenter le taux de refus ?).

27 Voir notamment Sophie Barluet, *Les revues françaises aujourd'hui : entre désir et dérives, une identité à retrouver. Rapport de mission pour le Centre National du Livre*, avril 2006, <http://www.centrenationaldulivre.fr/?Les-revues-francaises-aujourd-hui> ; sur la procédure actuelle de subvention, <http://www.centrenationaldulivre.fr/?SUBVENTIONS-AUX-REVUES> ; pour la composition de la commission chargée des sciences humaines et sociales (qui traite aussi de subventions à des livres), <http://www.centrenationaldulivre.fr/?Commission-sciences-de-l-homme-et>.

pour attribuer à la fois un label et une part de ressources rares, se sont ajoutés ces dernières années les portails de revues en ligne, francophones (Erudit, Cairn, Persée, Revues.org) ou internationaux (Jstor, Muse...), qui, lorsqu'ils ne sont pas exclusivement dédiés aux collections d'un éditeur ou d'un ensemble d'éditeurs, demandent aux revues une démarche de candidature avant, là aussi, de leur affecter une part de leurs ressources (personnel pour le suivi, heures travaillées pour la numérisation, espace serveur...). Comme pour le CNRS et le CNL, on peut observer que les procédures sont variées, souvent, quoique inégalement, claires quant aux personnes qui décident et aux critères généraux pris en compte, et peu contestées dans leur ensemble, moins encore dans leur existence même, par la communauté scientifique – même si telle décision précise, voire tel critère, peut à l'occasion faire l'objet d'un débat, il ne s'étend guère au-delà des personnes directement concernées.

L'expérience du Conseil scientifique du CLEO

Ne connaissant pas, ou indirectement (pour avoir « fait candidater » une revue) les autres lieux de sélection, je vais revenir ici sur mon expérience au Conseil scientifique du CLEO, dont la sélection des revues acceptées sur le portail Revues.org est une des tâches, à côté de la discussion de la politique générale du laboratoire. Cette structure me semble être l'héritage de ce qui avait été pensé comme une sorte de méta-comité de rédaction : comme un comité de revue sélectionne des articles tout en menant d'autres tâches (gestion des comptes rendus, parfois organisation de journées d'études, etc.) et en discutant de la ligné éditoriale générale, le Conseil scientifique sélectionne des revues (et plus récemment des collections de livres) tout en discutant par exemple du développement de Calenda ou d'Hypothèses, autres productions du CLEO, et plus généralement de la situation de l'édition électronique en France. Toujours comme dans une revue, le Conseil regroupe certains membres du laboratoire qui, concrètement, fabrique et entretient le portail, ainsi qu'une majorité de chercheurs, enseignants-chercheurs ou documentalistes en poste ailleurs, mais choisis pour leur intérêt pour l'édition électronique et leurs compétences dans des disciplines variées. Pendant les premières années, ce groupe expertisait directement les revues proposant leur candidature. Cependant, toutes les disciplines n'étaient pas représentées, la présence d'un-e historien-ne par exemple ne garantissait pas forcément une capacité à expertiser toute revue d'histoire, et toutes les revues ne se définissaient pas par une discipline (ainsi, beaucoup sont spécialisées par aire géographique ou linguistique) : d'où le recours rapidement mis en place à des experts extérieurs.

Celui-ci est également dû au fait que regarder de près une revue (ou un projet de création de revue, dans la mesure où nous demandons au moins un numéro complet pour expertiser de tels projets) prend du temps. Or, contrairement sans doute à ce que beaucoup d'entre nous avaient anticipé, il est souvent nécessaire de regarder de très près une revue pour décider s'il s'agit du type d'objet éligible, défini pourtant de façon *a priori* très ouverte comme « des revues scientifiques de sciences humaines et sociales et qui souhaitent participer à la diffusion des savoirs »²⁸. Si le dernier membre de phrase se réfère à des critères propres au portail et faciles à objectiver dans un formulaire (les revues doivent mettre à disposition au moins une partie de leurs articles en libre accès, se former à l'édition électronique et assurer elles-mêmes certaines tâches), il n'en va pas de même de la notion de « sciences humaines et sociales ». Il a fallu décider au cas par cas, tout en essayant de construire une jurisprudence cohérente en elle-même et avec la politique scientifique générale du portail, par exemple sur des revues touchant à l'environnement, la préhistoire ou les sciences cognitives et qui semblaient ne relever qu'en partie des sciences humaines et sociales. Les décisions ont finalement été très inclusives, d'autant qu'il n'existe pas d'outil éditorial exactement équivalent à Revues.org pour les autres disciplines. En partie pour des raisons voisines, c'est-à-dire l'absence d'équivalent exact du portail hors de France, s'est posée la question de deux implicites :

28 Delphine Cavallo, art. cit.

celui de la nationalité des revues (si j'ose dire) et celui de leur langue de publication, loin d'être toujours corrélés. Le portail est aujourd'hui en cours d'internationalisation, à la suite de choix également inclusifs dans le passé, mais qui n'étaient pas toujours évidents en termes d'attribution de ressources.

Enfin, il a évidemment fallu décider de ce qu'était une revue scientifique. Aucune définition n'avait été fixée *a priori*, sans doute pour le meilleur, car il aurait été difficile d'anticiper la variété des situations rencontrées. Sans doute pourrait-on arguer que certains choix du passé (quoique en très petit nombre) ne correspondraient pas aux critères d'aujourd'hui, mais il me semble qu'un bon équilibre a été atteint entre cohérence d'ensemble (au sens d'évitement d'injustices manifestes dans le traitement des candidatures selon le moment de dépôt) et évolutions nécessaires. Ainsi la catégorie « cahiers » a-t-elle été créée, en essayant de ne pas instaurer de hiérarchie avec les « revues », pour accueillir essentiellement des productions plus locales (en termes de taille et de diversité du comité de rédaction) mais pourtant de bonne qualité, avec l'idée, du reste, que ce caractère local pourrait dans certains cas évoluer du fait même de la mise en ligne. Les « bulletins » sont également à part parce qu'ils ne contiennent pas ou peu d'articles scientifiques au sens classique, notamment en termes de longueur. Cela dit, des objets un peu particuliers restent considérés comme « revues », en partie pour ne pas multiplier les catégories : ainsi de la revue bibliographique *Abstracta Iranica* ou de publications comme *Agone* qui relèvent plus de la diffusion de la recherche, parce qu'elles proposent des formats un peu différents de l'article scientifique classique (lui-même en réalité très variable entre disciplines...), et adoptent une ligne éditoriale plus explicitement politique, tout en utilisant des procédures de sélection claires et en produisant un contenu qui nous a semblé être de qualité.

Je passe sur bien d'autres cas particuliers, mais je voudrais souligner à partir de cette expérience que la définition d'un « périmètre de scientificité » n'a rien de trivial, qu'elle prend beaucoup de temps, en particulier si l'on veut éviter d'imposer des critères uniformes et au contraire s'attacher à maintenir l'ergodiversité²⁹ qui s'exprime entre sous-disciplines, lieux de production, générations, etc. Ainsi, si s'est imposée au fil du temps l'idée qu'une revue (ou un cahier) devait expliciter des critères de sélection des articles, nous ne prescrivons pas ce que doivent être exactement ces critères, mais en jugeons au cas par cas ; et si, lorsque nous regardons des revues, nous avons tous peu ou prou en tête quelques caractéristiques implicites d'un article scientifique (avoir des notes ou une bibliographie finale, se référer à une enquête de terrain, à des statistiques ou à un corpus textuel, être long de plus d'une page...), nous préférons les mobiliser, et éventuellement les amender, dans une discussion collégiale et appuyée sur une expertise plutôt que de recourir à un formulaire « à cases » qui réifierait ce type de critères.

Comme je l'ai dit, il a assez vite paru nécessaire de recourir à des experts extérieurs pour lire plusieurs numéros de chaque revue et juger de leur qualité, et parallèlement d'élargir la palette des disciplines représentées dans le Conseil lui-même. En effet, trouver le bon expert pour chaque revue est en soi un travail qui demande une connaissance minimale du champ – surtout si l'on veut éviter, ce qui n'est pas toujours possible, que la proximité d'objets se traduise en soutien ou opposition de principe. Un formulaire est fourni aux experts sollicités – suite à la demande des premiers d'entre eux –, mais il reste peu contraignant. Finalement, le Conseil scientifique dispose de trois types d'éléments : l'avis général de l'expert, rédigé sous une forme libre ; sa réponse à quelques questions plus fermées sur le public visé, le degré d'ouverture de l'autorité (régional, national, international), la régularité de parution et la visibilité de la programmation future, l'affichage de règles de sélection, le travail de mise en forme visible ou non dans les articles ; et des informations données par la revue lors de sa candidature : le statut juridique de la revue, son équipe éditoriale, son budget, sa durée

29 Ou diversité des manières de travailler, selon la belle formule de Sophie Roux, art. cit. Sur la notion voisine de respect des « cultures disciplinaires » comme aspect essentiel de l'évaluation collégiale, voir Michèle Lamont, op. cit.

d'existence, sa diffusion, ses projets d'édition électronique. L'expert-e est libre d'argumenter comme il ou elle le souhaite sur la qualité scientifique de la revue et l'intérêt de sa mise en ligne, sur la base de la lecture d'un nombre suffisant de numéros, qui lui sont fournis. Le Conseil scientifique doit donc aussi, *in fine*, comprendre comment l'expert a travaillé, s'il ou elle avait une conception plus ou moins restrictive de la qualité scientifique, etc. – de la même façon que les membres de comité de rédaction savent bien que certains relecteurs d'articles sont plus sévères que d'autres. Cela signifie que si l'avis de l'expert est d'un grand poids, il ne lie pas le Conseil scientifique, qui a déjà, dans le doute, recouru à une lecture directe et approfondie de la revue ou à une contre-expertise pour trancher. Les cas les plus complexes donnent en outre lieu à des discussions avec les personnels du CLEO ayant rencontré des représentants de la revue, puis avec la revue elle-même si nécessaire. De ce fait, on ne peut pas exclure un effet en retour des critères du Conseil scientifique sur les pratiques de certaines revues, même s'il me semble se restreindre à des points assez consensuels, comme l'affichage de critères de sélection ou l'existence d'une relecture orthotypographique des articles.

Ce qu'il faut à mon sens surtout retenir de cette procédure, dont nous souhaitons qu'elle permette de « construire une objectivité collégiale »³⁰, c'est qu'elle prend du temps, même si celui-ci est devenu difficilement mesurable depuis que discussions et votes se passent pour l'essentiel en ligne (les réunions étant plutôt réservées aux débats de politique générale, même si elles permettent aussi d'évoquer des cas complexes de revues ou de discuter de la création de catégories comme celles de « cahiers »). Une douzaine de candidatures de revues sont examinées tous les mois ou tous les mois et demi ; cela prend au moins quelques heures à l'expert et à certains membres de l'équipe du CLEO, au moins une heure à chaque membre du Conseil scientifique – alors même que certains cas « évidents » sont très vite tranchés. En 2008 comme en 2009, près de 600 messages par an ont été échangés entre membres du Conseil. Cette charge de travail explique que nous « faisons confiance » à l'examen préalablement effectué par d'autres pairs pour accepter certaines revues, principalement celles qui bénéficient du label CNRS, sans expertise, afin de nous concentrer sur les cas moins évidents. Elle explique sans doute aussi que plusieurs d'entre nous considèrent que, si ce type de travail a un sens lorsqu'il s'agit de répartir les ressources limitées de l'équipe du CLEO et de maintenir la qualité générale du portail, qui rejaillit sur chaque revue, il serait difficile, voire contre-productif, de le dupliquer à l'échelle de l'ensemble des revues françaises, voire mondiales, pour établir un « périmètre de scientificité » en soi, sans finalité précise. On peut ainsi noter que nous examinons des revues à l'entrée sur le portail et que, d'une certaine façon, nous faisons confiance à l'équipe du CLEO et aux lecteurs pour nous alerter si une revue, par la suite, « n'était plus scientifique » ; mais il n'est pas question d'actualisation en temps réel d'une liste – un des problèmes qui se poserait si l'on cherchait à établir un « périmètre de scientificité » valable à chaque instant.

L'expérience des listes de l'ESF et de l'AERES

Je ne reviendrai pas en détail sur les listes de l'AERES et les débats qu'elles ont occasionnés en France ; moins encore sur ceux qui concernent les listes de l'ESF, dites ERIH³¹. Je veux seulement évoquer quelques étapes-clés dans la constitution et l'évolution de ces listes, en rapport avec mes questionnements sur le comment et le pourquoi des classements de revues.

Les listes de l'ESF ont été mises en ligne avec une description très minimaliste de la façon dont elles avaient été constituées : quelques noms composant un comité d'experts pour chaque discipline (pour des listes de centaines, voire de milliers de titres en toutes langues), dont rien ne disait

30 Selon l'expression de Nicolas Dodier, art. cit., qui souligne notamment qu'objectivité et collégialité ne s'opposent pas, la première n'étant pas nécessairement associée, par exemple, à l'usage de formulaires. Pour une autre exploration de ce que peut signifier la collégialité en pratique, voir Michèle Lamont, *op. cit.*

31 European Reference Index for the Humanities : voir <<http://www.esf.org/research-areas/humanities/erih-european-reference-index-for-the-humanities.html>>. Comme dans plusieurs autres cas déjà évoqués, l'objectif explicite est de revaloriser les humanités, qui seraient désavantagées par rapport aux autres sciences par le manque d'instruments de mesure.

comment ils avaient été choisis, et une indication peu claire des catégories mobilisées : « top-quality », mais ensuite une hiérarchisation en A, B, C semblant renvoyer plutôt à un degré de diffusion (national/international), de « rayonnement » (notion plus imprécise pouvant évoquer les abonnements, les citations, la notoriété...) ou de spécialisation du domaine traité. Diverses absurdités (revues inexistantes, doublons, revues oubliées car trop interdisciplinaires...) ou biais manifestes (envers des langues, sous-disciplines, écoles de pensée...) ont vite été notés ; l'organisme a demandé aux revues – qui, au départ, ne connaissant pas la démarche de classement, n'avaient évidemment pas envoyé de renseignements – de réagir à leur classement et de fournir quelques informations de base (chiffrées : nombres d'abonnés, taux de refus...) dans la perspective d'une révision. La page web actuelle évoque, ce qui n'était pas le cas au départ, d'autres sources d'informations utilisées par les *panels pour classer* : « 15 expert panels sift and aggregate input received from funding agencies, subject associations and specialist research centres across the continent. » En même temps, elle annonce l'ambition de classer aussi les livres et les formats de publication « non-traditionnels ».

Ce classement se caractérise donc par l'opacité des informations reçues, des critères et des procédures utilisées pour en juger ; il n'a rien de bibliométrique, mais peut tout aussi difficilement être conçu comme reposant sur la « construction d'une objectivité collégiale », et cela avant tout parce que la tâche assignée semble impossible à réaliser de façon collégiale et délibérative : il s'agit en effet de classer toutes les revues du monde dans une discipline sur une échelle unique, sans objectif pratique bien déterminé, mais pour « servir de référence » dans l'absolu – et surtout pas, est-il parfois précisé en pure perte, d'utiliser les listes dans le cadre d'un recrutement. Bien que le classement ne repose pas sur la mesure de quelque chose de précis, il a en revanche, comme hiérarchie, déjà produit des effets – du découragement de certaines équipes à des débats au Conseil scientifique du CLEO sur la pertinence de ne pas expertiser les revues classées (ou classées A, ou classées A et B) par l'ERIH, en passant sans doute par des décisions locales plus lourdes de conséquences en termes de financements.

Si, en France, ce classement ERIH n'a pas été directement utilisé à grande échelle dans un but d'évaluation, c'est que l'AERES a publié peu après ses propres listes³². Dans certaines disciplines, comme l'histoire, aucune procédure ni aucun nom d'expert n'était donné, et il était manifeste que la liste de l'ERIH avait été reprise – en réalité avec des adaptations de détail et l'ajout d'un grand nombre de revues de provenance non explicitée. Ailleurs, l'ergodiversité régnait, dans les façons de concevoir les listes et dans ce qu'elles disaient de la définition de l'activité scientifique : on retrouvait plus ou moins de commentaires ou de détails sur les supports de publication autres que les revues, en droit par exemple (où des points étaient par ailleurs attribués à des critères variés, de l'ancienneté ou la diffusion sur Internet à la transversalité ou la diversité bibliographique), plus ou moins d'implication de représentants élus de la discipline (issus du Comité national du CNRS ou du Conseil national des Universités), ou encore plus ou moins d'interventions hiérarchiques, la liste pour la démographie étant cautionnée par des responsables de l'Institut national d'études démographiques.

L'intérêt de ce panorama de pratiques disciplinaires diverses et de leur jeu autour de la consigne (établir une liste hiérarchisée) trouvait toutefois sa limite dans l'objectif assigné à la liste : permettre de compter, dans les laboratoires que l'AERES avait pour mission principale d'évaluer, les membres « publiants » ou non, suivant leur nombre de publications d'une part d'articles dans des revues classées en A ou B, d'autre part de livres (l'ambition étant là aussi de proposer, mais dans un avenir indéfini, une liste hiérarchisée d'éditeurs). D'une certaine façon, avec un objectif aussi mécanique, l'adoption d'un critère bibliométrique arbitraire pour le choix des revues aurait été plus cohérent, sinon plus juste. Si, dans certaines disciplines, un travail réel de construction collective d'un

32 Le degré d'archivage des listes d'origine (été 2008) et intermédiaires sur le site de l'AERES est difficile à mesurer ; la page <<http://www.aeres-evaluation.fr/spip.php?page=recherche&recherche=revues>> permet toutefois de localiser à la fois les listes actuelles et au moins une partie des premières versions.

jugement des pairs a sans doute été réalisé, il s'est heurté, même si c'est dans une moindre mesure qu'à l'ESF, à l'ampleur de la tâche, au fait qu'aucune information n'était demandée aux revues (on imagine que les experts ont travaillé à partir de numéros, de sites web et d'idées générales sur leur réputation) et à la difficulté de trancher entre échelles de valeur : diffusion, caractère généraliste, international, vertu procédurale des modes de sélection d'articles, et toutes les versions possibles de la qualité intrinsèque des articles...

Après un tollé assez général, nombre de listes ont connu une seconde version – tandis que pour certains domaines, comme la littérature, l'AERES renonçait à constituer une liste. Là encore, les modalités de révision semblent avoir été extrêmement variées. En histoire, il s'est agi d'abord de changements à la marge, interprétés dans les couloirs, certes sans preuve tangible, comme les résultats de coups de téléphone directs des revues concernées. La liste compte aujourd'hui 1 700 titres environ, dans toutes les langues et sans indication claire de mode de constitution ; le principal changement est l'élimination de la hiérarchie en A, B, C, supposée mener à la constitution du « périmètre de scientificité ». Ce changement est commun à beaucoup de disciplines, mais pas à toutes. En sociologie-démographie, la liste est beaucoup plus courte, les noms des experts et leurs critères sont détaillés – et conduiraient sans doute à l'exclusion de nombre de revues présentes sur la liste s'ils étaient appliqués en histoire. Ces différences entre disciplines montrent que ce n'est pas parce que l'on s'en tient à un « périmètre » que l'on est d'accord sur sa largeur : toutes les revues un minimum scientifiques, ou bien les plus pérennes et procéduralement vertueuses ? Dernier exemple – mais la lectrice gagnera à se plonger directement dans la source pour se faire une opinion – : le préambule de la liste « anthropologie ethnologie », qui évoque un cas particulier : « Le champ de l'anthropologie biologique présente une spécificité. La commission considère que les experts chargés de l'évaluation des Unités de recherche travaillant totalement ou partiellement dans ce domaine, peuvent prendre en compte les revues ne figurant pas encore dans la liste de l'AERES mais présentant à leurs yeux toutes garanties de qualité scientifique. »³³

C'est dire clairement qu'il n'est pas possible, à l'échelle d'une discipline – et de toutes les revues du monde ! –, de se prononcer sérieusement de façon collégiale, même en termes de périmètre de scientificité, sur toutes les spécialités. Plus exactement, si le comité mobilisé par l'AERES devait directement trancher sur tel ou tel laboratoire, il y parviendrait sans doute, probablement en consultant des experts extérieurs. Mais ce qui lui est demandé – produire un outil pouvant ensuite être appliqué mécaniquement par un non-spécialiste – ne lui semble pas tenable dans ce contexte. Les difficultés inédites, et encore difficiles à caractériser, créées par ce type de classement non finalisé, ou plutôt indéfiniment finalisé, me semblent avoir été bien résumées par Nicolas Dodier : « Je crois que vous avez tout à fait raison pour distinguer des évaluations décisionnaires et des évaluations qui créent des notes qui ensuite flottent en attente de décision. Je crois que la création de l'AERES joue sur ces deux niveaux, puisqu'elle crée des notes qui ensuite sont susceptibles d'être utilisées dans des contextes très divers qui vont, au fond, les réinjecter dans leurs propres évaluations »³⁴.

*

Je ne prétends pas tirer de conclusions définitives de ce parcours cavalier au travers de pratiques finalement encore nouvelles – mais qui me semblent avoir été adoptées, finalement, par la communauté scientifique avec une rapidité étonnante, ce qui devrait nous poser des questions sur l'attrait des classements auprès des anciens bons élèves que nous sommes en général³⁵. J'en retiens pour ma part qu'il est important de sortir de discours un peu trop simples assimilant par exemple bibliométrie et classement, intervention de collègues et construction d'un véritable jugement des

33 <http://www.aeres-evaluation.fr/IMG/pdf/Liste_revues_anthropo_ethno_nov09.pdf>

34 Intervention à la 4^e séance du séminaire « Politique des sciences », décembre 2009, retranscrite sur <<http://pds.hypotheses.org/373>>.

35 Je renvoie à l'intervention de Cédric Lomba et Sophie Pochic dans la présente journée pour une proposition plus sociologique, à partir du cas des ingénieurs, sur cet élément ici donné un peu à la légère, et à celle de Philippe Büttgen pour une discussion des attraits et des périls de l'auto-évaluation.

pairs ; et que la critique nécessaire de la construction de tel ou tel indicateur ne doit pas trop détourner notre regard de la prise en compte du but dans lequel il est construit, de ses usages (par qui, et pour quoi faire) et de ses effets, l'objet mesuré ne restant pas statique ni même passif.

Quant à la question plus précise du classement des revues, il me semble que l'expérience française des listes AERES montre bien que toute la bonne volonté de ceux qui acceptent de s'impliquer dans l'amélioration de telles listes ne peut rien si leur but final reste de produire un outil applicable sans connaissance du domaine, et non pas des informations supplémentaires pour des pairs décidant collégialement et suivant des critères plus variés et évolutifs. Elle montre aussi que la constitution de telles listes est chronophage – sans parler de leur actualisation, si l'on ne veut pas qu'elles figent le paysage – et ne peut se faire sérieusement (avec un réel examen des revues) que si l'on envisage un « périmètre de scientificité » relativement réduit. Quel est alors l'apport pour la communauté, en dehors de toute visée officielle d'évaluation ou de pilotage, d'une telle liste constituée à grands frais ? Elle ne pourra pas résoudre toutes les questions de type « où publier ? » ou « comment juger de ce dossier de publications ? », parce que, comme dans le cas de l'anthropologie biologique évoqué plus haut, ceux qui auront constitué la liste n'auront pas pu prendre en compte toutes les langues, spécialités, etc. Surtout, quand bien même elle aiderait à la décision dans de tels cas, il peut toujours paraître plus simple, comme le font actuellement, mais informellement, nombre de membres de jurys de recrutement en France, de recourir à une rapide recherche en ligne : que la revue ait elle-même ou non un site web, cela permet en général (à partir de portails, de sites de bibliothèques, de sites personnels d'auteurs ou membres de comité de lecture, de mentions d'une activité de relecteur...) de se faire rapidement une idée de ceux qui la font, de ce qu'elle contient, voire de ses procédures, de sa diffusion ou de sa réputation, à partir de critères multiples et qui peuvent être remodelés au gré de la discussion. De plus en plus d'informations qui peuvent aider à mesurer, de façons diverses, les différentes formes de « grandeur » d'une revue, sont librement accessibles, jusqu'aux statistiques de fréquentation ou de téléchargement sur certains sites : peut-être cette profusion, qui serait alors à encourager (non par la production de nouveaux indicateurs, mais par la publicisation de ceux qui existent, notamment par les revues elles-mêmes), peut-elle permettre, au moins à la marge, de lutter contre la réduction à une échelle unique.

Par ailleurs, on pourrait souhaiter qu'une telle démarche soit surtout celle d'auteurs qui cherchent où publier, tandis que les recruteurs et autres évaluateurs pourraient tout simplement lire les publications elles-mêmes : non pas toutes, bien sûr, mais une sélection, de préférence effectuée et justifiée par le candidat lui-même, comme cela se pratique dans nombre d'institutions à l'étranger, voire en France. Il me semble facile de choisir entre allouer du temps à une telle démarche, ou bien à l'élaboration de listes générales à appliquer ensuite mécaniquement³⁶. Bien évidemment, cette position reste aujourd'hui en partie utopique face à l'imposition de procédures choisies pour des raisons tout à fait extérieures à la communauté scientifique. Toutefois, *a minima*, cette communauté peut choisir de défendre elle-même le jugement des pairs comme étant autre chose que le localisme ou l'arbitraire et de ne pas constituer elle-même les instruments de son contrôle par l'extérieur.

36 Ce que formule en termes plus choisis Sylvain Piron, « Quelques raisons de refuser tout classement des revues », décembre 2008, <http://evaluation.hypotheses.org/275> : « Soyons attentifs à la seule motivation de cette surveillance administrative de la liberté intellectuelle : il s'agit de “définir l'ensemble des revues dont les articles n'auront pas à être réévalués par les évaluateurs”. Et nous abdiquerions une si grande liberté pour un prétexte si médiocre ? »