Plaidoyer pour une bibliométrie non occidentale.

Pansu, Pascal, Nicole Dubois et Jean-Léon Beauvois. 2013. Dis-moi qui te cite, et je saurai ce que tu vaux. Que mesure vraiment la bibliométrie ? Grenoble : Presses universitaires de Grenoble.

Par Hervé Regnauld



L'évaluation des scientifiques est un sujet débattu et clivant. D'un côté, certains scientifiques sont farouchement opposés à toute forme d'évaluation récurrente, systématique et obligatoire, de l'autre, des instances revendiquent le droit de contrôler le travail des scientifiques qu'elles salarient et donc d'évaluer la qualité de leur production scientifique. La première position (refus systématique) est illustrée par divers ouvrages dont celui, récent d'Alain Abelhauser (2011). Elle reproche fondamentalement à toute évaluation d'être folle au sens propre, c'est-à-dire d'empêcher la réflexion rationnelle. En effet, l'évaluation, quelle qu'elle soit, commence par trier, ensuite elle hiérarchise, et enfin elle élimine. Le tri correspond à la démarche qui veut qu'on évalue les articles de telle revue et pas les chapitres de tel ouvrage, ou les publications chez tel éditeur et pas chez tel autre, ou les travaux en telle langue et pas en telle autre. Il y a donc tri arbitraire. La hiérarchie est établie quand des critères, quels qu'ils soient, servent à déterminer qu'un article est meilleur qu'un autre. Le critère le plus utilisé est le nombre de fois où l'article est cité. L'élimination est acquise quand les articles dits « pas bons » ne méritent plus d'être lus. Il y a donc folie au sens propre : un

jugement d'exclusion est porté à partir d'un échantillonnage incomplet et selon des critères qui imposent la différenciation, excluent l'égalité, donc l'altérité. Le respect de l'altérité passe par l'acceptation comme égal d'un objet qui est différent. La folie commence quand on ne sait plus qui est l'autre.

En faveur de l'évaluation, il y a un argument politique de fond : l'État, qui paye les scientifiques, doit s'assurer que ces scientifiques remplissent effectivement les missions que leur statut leur impose. L'évaluation est donc un acte démocratique qui a pour but d'établir la transparence du rapport entre rémunérations et missions pour tous les agents de l'État. Le livre de Pascal Pansu, Nicole Dubois et Jean-Léon Beauvois, *Dis-moi qui te cite, et je saurai ce que tu vaux. Que mesure vraiment la bibliométrie ?*, défend ce point de vue politique et ajoute un deuxième argument, tout aussi pertinent : un agent de l'État peut ne pas occuper son poste toute sa vie et doit donc pouvoir être remplacé. Il doit donc penser son travail sous la forme d'une sociabilité non individuelle.

[…] cette sociabilité est catégorielle, en ceci qu'elle conduit l'individu à se vivre, en tant qu'agent social, comme élément d'une catégorie […] Elle est aussi organisationnelle en ceci qu'elle détermine l'entrée de l'individu dans ces artefacts sociaux que sont les organisations. C'est cette forme de sociabilité liée au pouvoir social qui génère des identités sociales […] qui aujourd'hui contraste avec l'individualité. (p. 12)

L'évaluation est donc ce qui permet à des agents sociaux de se situer les uns par rapport aux autres dans le cadre organisationnel qui leur est commun. Dès lors que les résultats de cette évaluation sont publics (en respectant les données confidentielles), le fonctionnement de l'organisation devient transparent, donc plus démocratique. Dans le cas particulier de l'Université, l'évaluation soulève un problème spécifique : peut-on, en tant qu'universitaire, refuser d'être évalué, mais accepter de siéger dans un jury de Doctorat, d'Habilitation à Diriger les Recherches ou d'être *referee* dans une revue ?

À partir de ce point favorable au principe de l'évaluation, de départ pro-évaluation, les auteurs tentent d'établir que, pour les scientifiques, l'évaluation bibliométrique est un bon outil, bien que les conditions actuelles de sa mise en

œuvre soient critiquables. Ils dénoncent donc ce qui doit être abandonné et proposent une modalité d'évaluation bibliométrique qui, selon eux, serait acceptable, juste et véritablement internationale.

Ils commencent par décrire précisément les diverses procédures bibliométriques (indices H et G, Impact factors (IF) etc.) et comparent les différentes valeurs obtenues selon qu'on utilise Google Scholar (open access), Scopus (chez Elsevier), Institute for Scientific Information (Thomson Reuters). Tous ces outils n'utilisent pas les mêmes bases de données (10 000 revues pour ISI, 19 000 pour Scopus) et ne trient pas les publications selon les mêmes critères. Certaines bases de données prennent en compte des cours mis en ligne (Google), d'autres pas… De cette analyse précise ressort un point fort : chaque base de données a sa part d'arbitraire. Les calculs mis en œuvre sont aussi légèrement différents d'un outil à l'autre et les IF ne sont pas calculés sur les mêmes laps de temps. Cela implique donc un choix implicite : si on utilise l'outil *x* plutôt que *y* ou *z*, on accepte une part d'ombre qui réside dans les mécanismes de calcul et la vitesse de leur actualisation. Et si on utilise et *x*, et *y* et *z*, on obtient trois valeurs différentes. À ce stade de la lecture du livre, le lecteur se demande pourquoi les auteurs tiennent encore à l'évaluation bibliométrique.

La réponse est fournie par un détour. Les auteurs, en effet, ne reprennent pas les arguments classiques (vus ci-dessus), mais apportent un nouvel élément, construit à partir d'une démarche qu'ils ont menée au cours des années précédentes. Ils ont envoyé un questionnaire à plusieurs centaines de scientifiques et ont reçu en retour 128 réponses. Ils précisent que leur échantillon n'est pas représentatif, mais qu'il a cependant une valeur informative. Les analyses statistiques (simples) des réponses sont très claires. Les scientifiques dits « durs » sont attentifs aux IF et citations, les scientifiques des Sciences Humaines et Sociales (SHS) n'y prêtent pas attention. En Science de la Vie, 83 % des interrogés signalent que trois ou plus de leurs derniers articles sont dans des revues à IF, en SHS c'est 25 %. Autrement dit, les trois quarts des publications en SHS échappent à l'IF, contre à peine 20 % en Science de la Vie. Ceci n'est pas une nouveauté. D'autres résultats confirment la nette différence entre les divers domaines scientifiques. Quand on demande : en combien d'années avez-vous publié vos trois derniers papiers (articles et chapitres

d'ouvrage) ? Ce sont les SHS qui ont les délais les plus longs (plus de 50 % des interrogés mettent plus de 3 ans pour trois publications), tandis qu'en Sciences de la Vie les délais sont très courts : 80 % des interrogés publiant 3 articles en une seule année. De façon assez surprenante cependant, tous les scientifiques déclarent choisir la revue dans laquelle ils publient en fonction de sa « réputation ». Il est donc logique de s'interroger sur la façon dont une réputation s'établit pour des sciences aussi différentes les unes des autres. Selon les auteurs, la notion de revue réputée, que tous les scientifiques acceptent, est la preuve que l'évaluation, même chez ceux qui la récusent, a un sens et une efficacité réelle. Quoiqu'on dise par ailleurs, le fait qu'on différencie les revues en plus ou moins réputées implique qu'on accepte le principe d'une hiérarchisation et qu'on en assume les conséquences en choisissant explicitement de soumettre les projets d'articles là et pas ailleurs. Il y aurait donc une évaluation implicite dans chaque stratégie de publication.

La suite de l'ouvrage prend acte de cet implicite et détaille les mécanismes qui permettent à une revue d'être considérée comme plus ou moins réputée. Dans une bibliothèque, la revue la plus réputée est celle qui est le plus souvent demandée et il est donc logique d'en avoir plusieurs exemplaires et de les ranger le plus près possible du comptoir de délivrance aux usagers. C'est au départ pour rationaliser l'occupation de l'espace dans une bibliothèque que la bibliométrie a été inventée. En tant que telle, c'est une science utile, solide et elle n'est pas condamnable. Il se trouve simplement que c'est, à la base, la demande qui fait la notoriété ; « Pourtant aucun épistémologue n'avancera l'audimat comme critère de validité et de valeur d'un écrit scientifique » (p. 79).

L'argument premier est donc clair : la bibliométrie mesure la notoriété, qui elle-même dépend des prescripteurs de notoriétés. Elle ne mesure pas la qualité des articles. Ceux qui élaborent la notoriété d'une revue sont ceux qui conseillent sa lecture. Directement, ce sont les enseignants qui disent aux étudiants de lire *x* ou *y*. Indirectement, ce sont les scientifiques qui, dans leurs propres articles, citent tel autre article dans telle revue. À l'évidence, les éditeurs commerciaux d'une revue ont tout intérêt à favoriser les citations de leur revue et privilégient forcément toutes les stratégies qui vont multiplier les citations. La plus commune est de publier un panel de revues aux thématiques non semblables, mais

voisines de façon à ce que les scientifiques de l'une citent l'autre. Une revue plus généraliste vient alors en chapeau surplombant pour publier des articles synthétiques et est à son tour citée par toutes les autres. Elle peut augmenter ses tarifs d'abonnement.

Le second argument est plus fort encore. Il dit que la bibliométrie telle qu'elle est utilisée pour évaluer les scientifiques est un instrument de mesure faux. Un instrument juste n'altère pas l'objet qu'il mesure. Un thermomètre ne fait pas varier la température de l'air. En revanche, si une revue est mesurée comme bonne (avec un gros IF), elle attire plus d'auteurs, elle peut se permettre de refuser plus d'articles, elles valorisent donc davantage ceux qu'elle publie et elle est citée plus souvent. Surtout, le scientifique qui publie chez elle est réputé meilleur…

Un troisième argument est politique. Les auteurs prennent l'exemple de la psychologie sociale dont ils sont spécialistes. Supposons, disent-ils, qu'un Cambodgien élabore un nouveau concept en psychologie sociale et qu'il soit 1) inspiré par les travaux d'un de ses collègues cambodgiens (concept publié dans une revue en cambodgien) et 2) validé par des protocoles expérimentaux mis en place sur tant de Cambodgiens. Le choix du Cambodge n'est pas anodin : les années 1975-1977 y sont hyper violentes et leurs conséquences sur la psychologie des survivants commencent à laisser voir des traces dont la fréquence et l'intensité peuvent permettre des généralisations dépassant les cas individuels. Mais comment faire passer dans une revue « internationale » un travail qui ne repose que sur des sources cambodgiennes ? Les *referees* vont demander que le travail s'inscrive dans une perspective plus globale (les génocides en général !) et qu'il cite les travaux de référence (en anglais) sur ce sujet. Ce qui est en jeu ici est alors la spécificité d'une psychologie cambodgienne face à la généralité d'une psychologie dite universelle. Les auteurs n'hésitent pas à dire qu'il faut « contester l'universalité de la psychologie dominante qui est aujourd'hui la psychologie occidentale » (p. 88). La psychologie occidentale est fondée sur un présupposé que « rien ne justifie sur un plan empirique » (p. 87). Ce présupposé établit que les études psychologiques faites sur la base individuelle, sur la base d'un moi, sont plus valides que celles qui sont faites sur la base de l'appartenance à un groupe. De

manière différente, la psychologie non occidentale ou psychologie indigène privilégie le groupe. Elle est « l'étude du comportement humain et des processus mentaux au sein d'un contexte culturel qui s'appuie sur les valeurs, les concepts, les systèmes de croyances, les méthodologies et autres ressources indigènes » (p. 89).

Les arguments scientifiques par lesquels sont jugés les articles dans les revues occidentales (souvent anglophones) ne sont donc pas universellement valides. Il n'existe tout simplement pas de critères universels de scientificité pour la psychologie sociale. Cela implique qu'il ne peut pas exister de critères permettant de hiérarchiser les revues.

À ce stade de la lecture, et pour la seconde fois, la bibliométrie est terriblement mise en cause : elle peut être détournée pour des raisons commerciales, elle déforme ce qu'elle mesure et elle véhicule un impérialisme culturel occidental. Pourquoi donc continuer à défendre une évaluation bibliométrique ?

Les trois auteurs de l'ouvrage affirment que l'évaluation est une nécessité et qu'elle ne doit pas être uniquement qualitative : « refuser la quantification en tant que telle revient à refuser la clarté de la pensée et/ou louanger le flou d'une pensée » (p. 15). Ils ajoutent : « Ce n'est ni l'idée d'impact, ni la quantification qui font aujourd'hui problème à nos yeux. Ce qui fait vraiment problème, c'est la façon dont cette idée a été exploitée et le contexte géopolitique de cette exploitation » (p. 108).

Ils cherchent donc à construire une forme de bibliométrie qui s'ajouterait à une évaluation qualitative et qui n'aurait pas les défauts signalés auparavant. Ils proposent donc d'abord une nouvelle façon de construire une base de données, puis une nouvelle liste de critères quant au fonctionnement des revues. Une commission internationale d'experts établirait une liste de — *a minima* — trente nations où l'on sait que se pratique une recherche scientifique dans la discipline, et les revues de toutes ces nations (donc en langues différentes) seraient toutes prises en compte, pour peu qu'elles fonctionnent selon les modalités suivantes : *peer-review* ; comité éditorial national (en langue indigène) ; comité de lecture international ; articles avec résumé étendu (1000 à 2500 mots) dans une autre langue que celle du texte principal ; mots-clés en plusieurs langues ; références

bibliographiques renvoyant à des revues de langues différentes. Les chapitres d'ouvrages sont également pris en compte avec les mêmes exigences. Si toutes ces conditions sont respectées, les trois auteurs pensent que « l'impact […] est de droit un excellent critère d'évaluation d'une revue » (p. 108).

Ce livre est donc une façon de mettre les pieds dans le plat dans le débat actuel sur l'évaluation. Il est peu courant que des scientifiques « non durs » soient partisans d'une évaluation bibliométrique et il est encore plus rare que ce soit en vue de dénoncer l'impérialisme scientifique occidental, pour promouvoir une approche favorable aux sciences replacées dans un contexte « indigène ». En dépit de ce louable engagement politique, l'ouvrage laisse largement inexplorée une face cachée de la bibliométrie : que faire quand un *referee* écrit que votre article est intéressant, mais qu'il oublie de discuter de l'avis de *x* ou de *y*, avis exprimé dans tel article de telle revue ? Que l'article soit écrit dans une langue ou une autre, que les *referees* parlent ou pas la langue des auteurs, peu importe ; dès lors que les citations sont prises en compte dans l'évaluation, il y aura toujours un *referee* pour suggérer que telles références enrichiraient votre propos et permettra à votre article d'être publié.

Le fond du problème de la bibliométrie est de mesurer quelque chose qui n'est pas le contenu scientifique d'un article, mais son inscription dans un réseau de scientifiques (je te cite, tu me cites…) en favorisant les disciplines à gros effectifs. Pourquoi, pour évaluer la qualité du travail d'un scientifique, ne pas, tout simplement, lire son travail indépendamment du support ?

Résumé

L'évaluation de la science est en fait l'évaluation des scientifiques et de ce fait elle est radicalement récusée par nombre d'entre eux au nom des libertés académiques. D'autres scientifiques l'acceptent — les auteurs de ce livre, par exemple — mais ils proposent qu'elle fonctionne sur d'autres bases que celles qui existent actuellement. Ils insistent en particulier sur le fait que la science est un produit culturel s'inscrivant dans divers contextes, pas tous occidentaux, et que les environnements scientifiques non occidentaux sont aussi valables que les autres.

Bibliographie

Abelhauser, Alain, Roland Gori, Marie-Jean Sauret. 2011. *La folie évaluation. Les nouvelles fabriques de la servitude*. Paris : Éditions Mille et une nuits.