

La bibliométrie évaluative : une redéfinition des valeurs scientifiques.

Rémi Goasdoué

Université Paris-Descartes, Laboratoire EDA

remi.goasdoue@parisdescartes.fr

Résumé

La bibliométrie évaluative est abondamment critiquée pour ses méthodes ou ses usages, mais plus rarement pour ses effets. Plus précisément la question de ses effets est traitée principalement à partir de cas caricaturaux allant des jeux stratégiques à la fraude. L'ambition principale de cette synthèse est de défendre l'intérêt d'une analyse d'influences moins visibles et non moins importantes de la pression à la publication sur les pratiques de recherche et les écrits scientifiques. L'analyse des effets du *publish or perish* sur, le choix des sujets de recherche, la manière de poser les problèmes, de traiter les données et d'écrire, montre une forme de participation/adhésion implicite et parfois involontaire des chercheurs à cette culture. L'influence croissante des considérations quantitatives dans l'évaluation de la recherche en dépit des critiques nombreuses et légitimes de la bibliométrie témoigne probablement de l'existence de ces effets plus profonds. En transformant la citation en unité de compte, en monnaie, la bibliométrie conduit à une redéfinition de la valeur scientifique et à des oppositions éthiques dans le champ scientifique, comparables à celles qui naguère opposaient les valeurs « aristocratiques » et « marchandes ».

Mots-clés

Bibliométrie évaluative, valeurs scientifiques, scientométrie, publications scientifiques, évaluation de la recherche

Pour citer cet article : Goasdoué, R. (2015). La bibliométrie évaluative : une redéfinition des valeurs scientifiques. *Evaluer. Journal international de Recherche en Education et Formation*, 1(1), pp. 45-64.

1. Introduction : une critique de la critique

La bibliométrie évaluative est abondamment critiquée pour ses méthodes ou ses usages, mais plus rarement pour ses effets. Le premier objectif de ce texte est de dépasser la dénonciation de la quantophrénie et de décrire des effets plus discrets sur les productions scientifiques que les exemples caricaturaux de manipulations

stratégiques qui émaillent les discours critiques. Bourdieu (1996) affirmait à propos de la télévision qu'il serait réducteur d'expliquer les choix journalistiques en se référant uniquement aux sources de financement des médias. Bien que légitime dans certains cas « caricaturaux » de censure étatique ou commerciale, cette critique ne rend pas compte des multiples choix conscients ou inconscients qui s'opèrent quotidiennement dans le champ

journalistique. La transposition de cette « critique de la critique » au champ scientifique et aux effets de la bibliométrie est le point de départ de ce travail.

Raisonnement en termes d'effets pourrait conduire à privilégier des exemples patents comme des pratiques déviantes allant des jeux stratégiques à la fraude. Or suivre la proposition de Bourdieu invite à comprendre les choix plus courants que font les chercheurs en termes de sujets de recherche, de méthodes d'enquêtes, de traitements des données, de références bibliographiques. Les déterminations multiples de ces choix posent le problème de parvenir à cerner ce qui tient à l'évolution des publications, au *zeitgeist* ou plus spécifiquement à la bibliométrie évaluative. L'inscription historique de ces évolutions conduit à modérer l'emploi de raisonnements de type « cause à effets ». Ceci amène donc à reconsidérer la proposition de départ : l'enjeu de cette synthèse n'est pas de décrire des « effets » dans un schème causal, mais de cerner dans l'évolution des pratiques de publication concomitantes à l'apparition de la bibliométrie évaluative. Cette formulation sous-entend implicitement que la bibliométrie est en partie extérieure au champ scientifique. C'est le cas historiquement car, jusque dans les années 1970, l'évaluation de la recherche était confinée aux dialogues entre chercheurs dans le processus de relecture des articles. Cependant, l'autonomie des sciences, lentement conquise à travers notamment le processus d'évaluation par les pairs, est en partie remise en cause par l'émergence des divers procédés de quantification issus de la scientométrie. Par les conventions sur lesquelles ils reposent, leur mode de calcul et plus encore leurs usages, ils offrent en effet une prise sans précédent aux acteurs externes au champ scientifique pour en régir le fonctionnement.

Gouverner par les nombres n'est pas une idée neuve, mais l'usage généralisé d'indicateurs, qui deviennent progressivement des objectifs, est toujours une source importante d'hétéronomie. Le second objectif de cette synthèse est de décrire les évolutions des modes de régulation du champ scientifique induites par l'utilisation de la bibliométrie. Toute quantification sur des comportements humains a nécessairement en retour un effet sur ces comportements et, probablement d'autant plus grand, qu'elle est omniprésente dans beaucoup d'outils de recherche documentaire, voire sur les sites de chercheurs qui assortissent la présentation de leurs travaux de mesures d'impact des revues dans lesquelles ils sont publiés.

L'implication des chercheurs *volens nolens* à cette quantophrénie n'est pas clairement abordée dans les deux types de critiques de la bibliométrie. Beaucoup de discours font de la bibliométrie un tout homogène extérieur au monde scientifique. Cependant, la bibliométrie n'est possible qu'à certaines conditions qui sont produites par les chercheurs eux-mêmes. Elle repose en effet sur une normalisation de la production scientifique sans laquelle produire les principaux indicateurs de bibliométrie est un travail très coûteux. L'ISI, compagnie fondatrice du Science Citation Index (SCI) a frôlé la faillite à ses débuts en constituant des bases à partir de données hétérogènes et non informatisées. La poursuite du projet de base n'a été possible que grâce aux financements de la National Science Foundation (Wouters, 2006). La viabilité financière du projet bibliométrique a été assurée ultérieurement par l'arrivée des bouquets de revues numériques et la hausse considérable des coûts qui l'a accompagnée. Ainsi, le facteur d'impact, initialement conçu pour la gestion des abonnements dans les bibliothèques

universitaires, est devenu une arme de négociation des abonnements car, sous le double effet de la numérisation et des concentrations/rachats des éditeurs scientifiques et des revues, en dépit des gains de productivité liés à l'informatique et au travail d'édition produit par les scientifiques, les coûts à la page ont augmenté de 250% à 580 % selon les disciplines (Carbone, 2010).

Si la bibliométrie était seulement confinée à la régulation des marchés de l'édition scientifique, l'essor considérable depuis une dizaine d'année des revues en accès libre devrait conduire à un désintérêt progressif pour ces statistiques d'audience. Or c'est loin d'être le cas comme en témoigne la mise au point d'applications pour *smartphones* qui permettent certes d'être informé de nouvelles publications indexées par des mots-clés mais aussi de recevoir des « alertes » à toute nouvelle citation de ses propres travaux. Il serait donc réducteur de ne voir dans la bibliométrie qu'un outil de gestion documentaire ou de négociation du montant des abonnements aux bouquets de revues. L'émergence de son influence s'inscrit dans un double mouvement : technologique, qui la rend possible, et social qui promeut la quantification au rang d'argument ultime ou de nécessité quotidienne. Cette combinaison entre la disponibilité des données (informatisées et publiques) et l'imprégnation, quelle que soit la sphère de la société, d'une « mentalité audimat » (Bourdieu 1996), explique, qu'en dépit de critiques fondées et réitérées, les usages évaluatifs de la bibliométrie persistent. Comprendre les ressorts de l'adhésion plus ou moins explicite des chercheurs est également un enjeu important pour cerner l'influence de la bibliométrie.

Ces questions seront abordées à travers une synthèse de travaux portant principalement sur des recherches en

sciences humaines. Ils seront présentés conformément à une structure canonique d'écrits scientifiques, de l'introduction à la bibliographie, au risque de disjoindre artificiellement des moments du travail de recherche, tout comme l'imposent les structures d'écrits scientifiques.

2. Le choix des objets de recherche : entre *zeitgeist* et quête du scoop

Choisir un objet de recherche est de longue date le premier acte pour être « visible » scientifiquement. Les stratégies sont multiples et anciennes : Matalon et Lemaine (1969) soulignent l'importance, dans le système de récompense du champ scientifique, de la quête de l'originalité qui est associée à celle de la primauté et en conséquence de la visibilité. La compétition pour la nouveauté est un moyen d'accéder à une forme de visibilité/reconnaissance pour les chercheurs, et donc un puissant moteur pour les découvertes. L'accès à une quantification de cette audience peut conduire à ne plus faire de paris audacieux d'originalité pas toujours payants en termes de visibilité. S'imposent au contraire des choix « maximisateurs » de visibilité, combinant les formes antérieurement connues mais optimisées par les gains d'audience.

2.1 La recherche d'une niche

S'insérer dans un champ balisé avec un « avantage compétitif » est à la fois une stratégie marketing et un schème fondamental d'entrée dans un champ de recherche que l'on trouve explicité dans les introductions d'articles du type : « De nombreux travaux ont montré que... mais peu se sont intéressés à... ». Ce positionnement, qui a en soi une pertinence scientifique, a également une influence sur le nombre de citations. Dans une étude dans le champ de la médecine d'urgence, Callahan *et al.* (2002) ont comparé des données bibliométriques de citations et de

jugements précédemment portés par des pairs lorsque ces recherches avaient été soumises à un colloque. Au moyen d'analyses de régression, ils montrent que le critère déterminant pour prédire le nombre de citations est d'abord le facteur d'impact de la revue, puis la taille des échantillons d'enquête, la nouveauté de l'objet de recherche et, pour une part négligeable, la qualité de la recherche (telle qu'évaluée lors de la présentation au colloque antérieur). En s'intéressant cette fois à la prédiction du facteur d'impact de la revue dans laquelle ces recherches sont publiées, le critère de nouveauté passe nettement devant celui de qualité et des autres critères méthodologiques comme l'échantillon, la formulation d'hypothèses, etc. Ceci souligne très clairement l'importance du choix du sujet, surtout pour être publié dans des revues à large diffusion. Tout se passe comme si, plus une revue est citée, plus elle aurait tendance à privilégier des sujets qui vont faire parler d'elle.

2.2 L'entrée par prise d'opposition

Outre cette première stratégie d'accession à la visibilité par la « nouveauté », il existe aussi celle du contre-pied de positions ou d'auteurs célèbres. Mehler et Bever ont ainsi fait une entrée fracassante dans le champ de la psychologie, en publiant dans la revue *Science* en 1967, une réplique-variante d'une tâche de conservation du nombre proposée par Piaget, montrant une acquisition bien plus précoce du dénombrement. Ne s'agissant pas d'une réplique exacte et donc pas précisément d'une réfutation, Piaget fit publier une réponse dans la même revue (son seul article dans la revue *Science*). La postérité ou plutôt les « post/contre-piagétiens » n'ont cependant retenu que le premier article. Ils sont du reste largement pourvoyeurs des citations ultérieures de l'article de Mehler et Bever. En effet, l'article de

Mehler et Bever a été cité 123 fois dans la base de données du Web of Science (WoS) et la réponse de Piaget 91 fois dans cette même base¹. Cet exemple est anecdotique dans la mesure où ces auteurs ont ultérieurement publié sur le langage et son acquisition délaissant les questions du développement du nombre traitées dans cet article : il eut pu en être autrement s'ils avaient été animés par la seule quête effrénée de citations. Bourdieu (2001) fustige les intellectuels qui s'autoalimentent d'oppositions factices, qui finissent par imposer comme légitimes des questions sans intérêt. Dans le cas précédent, il faut quand même remarquer l'efficacité de la technique car l'article de Mehler et Bever est de loin l'article le plus cité parmi les 43 publiés/recensés dans le WoS par Mehler, ainsi que pour Bever dont la base recense 11 articles. Ceci montre, en terme de citations, que tout contradictoire vit au dépend de celui qu'il critique !

2.3 Les sujets en vogue

Qu'il y ait des sujets porteurs n'a rien de neuf mais, les outils de documentation, internet et les réseaux sociaux produisent sur le champ scientifique les effets classiques de concentration des médias de masse. L'évolution des sujets à la mode est nécessairement plus rapide que lorsque les colloques annuels ou biannuels étaient le principal vecteur de l'actualité scientifique. Chercher une influence de la bibliométrie sur les choix des objets de recherche est pourtant difficile à montrer car c'est faire le pari de l'existence de choix opportunistes de certains chercheurs qui, par définition, ne peuvent être quantifiés. Pourtant, dans les « hottest articles » recensés par Scopus

¹ Ces résultats sont apparemment biaisés par le choix de cette base qui n'est pas spécialisée en psychologie ; cependant, on retrouve un rapport semblable entre le nombre de citations de ces deux articles dans Google Scholar (253 pour Mehler et Bever et 160 pour Piaget).

comme étant les 25 articles les plus cités, 8 en sciences sociales et 7 en psychologie portent sur Facebook. L'éternelle jeunesse des nouvelles technologies risque de ne pas reléguer de sitôt les recherches à ce sujet dans des revues confidentielles, si le critère de la nouveauté prime sur tous les autres.

La résonance entre émergence médiatique et débat scientifique peut également être abordée à travers l'exemple du programme PISA. Des travaux sociologiques sur la réception de ces résultats dans différents pays traitent de l'écho de ce programme sur les institutions d'éducation (Pons, 2011), mais une sociologie de la recherche sur ce sujet serait également très intéressante. Une enquête bibliométrique relative à l'impact de ce programme sur les revues scientifiques² (Dominguez et al., 2012) atteste évidemment de l'intérêt porté par les chercheurs à PISA. Cependant, limiter cette analyse au constat de l'augmentation du nombre d'articles publiés par an³, ou encore du nombre de numéros spéciaux qui lui est consacré, serait aussi réducteur qu'un décompte de citations. Dominguez *et al* (2012) entrent de manière plus qualitative dans la description de cette prolifération d'articles et font plusieurs constats intéressants pour décrire la brève évolution de ce champ de recherche en pleine émergence. Ils montrent que la plupart des articles sur la période 2002-2010 portent sur des aspects méthodologiques (78,6%) et peu d'entre eux reposent sur les données issues de PISA (12,1%), mais que cette proportion augmente considérablement avec le temps. Cette évolution montre d'abord que PISA est devenu dans le champ académique une source de données pour des recherches et, plus largement,

comment le milieu universitaire assimile, au sens plein du terme, une nouvelle source de données entérinant au passage une forme de délégation dans la production des données.

2.4 Les appels d'offre

La part croissante de financements sur projet joue aussi considérablement sur les choix de sujets de recherche. Cependant cette évolution est à double tranchant et n'est pas toujours le biais d'un management par les finances qui menacerait l'autonomie des universitaires. Les appels d'offre sont co-conçus par des scientifiques et l'évaluation des propositions est le plus souvent de leur fait également. Ainsi, toutes les catégories de jugements habituellement en cours dans les milieux scientifiques sont également à l'œuvre dans ces dispositifs. Cependant un raisonnement économique vient s'insérer dans les logiques scientifiques, il s'agit du postulat qu'un bon appel d'offre (qui finance l'excellence) a un taux de rejet de propositions élevé. C'est a priori que l'on retrouve intégré par les scientifiques dans la gestion des revues conduit à un certain nombre d'apories dues à l'oubli du caractère relatif de cette définition « économique » de l'excellence par la rareté. Si, comme c'est le cas, le nombre de publications par an croît considérablement, mécaniquement, à moins de créer un nombre considérable de nouvelles revues par an, le taux de rejet par an ne peut qu'augmenter sans qu'il s'accompagne nécessairement d'une amélioration des productions. De la même façon, les réorientations des financements récurrents vers des projets conduisent nécessairement à une augmentation des candidatures et donc du pourcentage de rejet. Le choix des thématiques des appels d'offre et les manières de répondre sont tributaires de ces jeux. Il doit être suffisamment large pour amener un nombre conséquent de

² Par l'utilisation des bases Eric, EBSCOhost, ISI Web of Knowledge

³ 4 articles par an en 2002, 30 en 2006, et 69 en 2010

propositions (en vue d'améliorer le taux de rejet et donc de légitimer les élus) et inversement les chercheurs jouent avec les thématiques pour faire financer leurs recherches comme le soulignent de nombreux témoignages :

« Le formatage des projets commence d'ailleurs au niveau doctoral : pour espérer un financement, nos jeunes collègues se trouvent maintenant obligés de proposer des recherches balisées, qui sortent aussi peu que possible des sentiers battus. Annonçant pratiquement leurs résultats à l'avance, ils tentent ainsi de garantir le retour sur investissement. S'il fallait évaluer les projets de la génération précédente sur de telles bases, la plupart seraient tout simplement refusés. Quelle ironie, quand on pense que ses représentants occupent aujourd'hui les postes d'évaluateurs. » (Gosselain, 2011, p. 133)

3. La production des données et leur traitement

La pression à la publication a des effets sur la production des données, le choix des données publiées, leur traitement statistique et les interprétations qui s'ensuivent. Seules quelques pistes seront évoquées ici, en allant des effets caricaturaux aux effets moins directement visibles.

3.1 La fabrique des données

Ce titre est plus pris au premier degré qu'au sens bachelardien de la construction des faits scientifiques, par certains chercheurs qui inventent des données. La fraude était, jusqu' il y a peu, l'affaire de quelques idéologues voulant à tout prix montrer l'efficacité ou l'importance d'un traitement ou la supériorité d'un groupe social. Ce qui n'intéressait que les historiens des sciences est devenu une préoccupation

institutionnelle face à la récurrence des fraudes (Fagot-Largeault, 2010). Cette préoccupation nouvelle pour l'inconduite scientifique a débuté dans les années 1980, date à laquelle la valorisation quantitative des publications a également fait son apparition. Sans que l'on puisse affirmer autre chose qu'une concomitance, nul n'est besoin de dépenser un million de dollars comme le suggérait un éditorial dans Nature (Marshall, 2000) pour estimer non pas l'évolution mais les causes des fraudes dans le champ scientifique. La précarisation des métiers de la recherche ou encore ses modes de financement ne sont en effet probablement pas plus étrangers à cette évolution que l'apparition du dopage dans les sports se professionnalisant rapidement. À défaut de données fiables sur la fraude, le nombre de rétractations croissant depuis une vingtaine d'années peut être interprété comme une efficacité accrue du travail des relecteurs (Fanelli, 2013). Pourtant, et au-delà des anecdotes qui alimentent les éditos des revues généralistes recensant des cas caractérisés de fraudes, voire de destruction de données par un concurrent dans un même laboratoire, une synthèse des enquêtes empiriques (par questionnaires anonymes) proposée par Fanelli (2009) recense des chiffres inquiétants, allant de 14% pour les fraudes considérées comme graves à 72% pour des inconduites moins explicitement identifiées comme de la fraude, par exemple l'exclusion de cas marginaux dans des échantillons⁴. Bien que ces effets ne puissent être attestés de manière catégorique, il demeure intéressant de constater qu'une intervention externe (d'incitation à la publication) en appelle une autre, celle d'agences institutionnelles de contrôle et

⁴ Ces chiffres sont ceux donnés lorsque l'on demande à des chercheurs s'ils ont constaté des fraudes, ils tombent à 2% et 34% lorsqu'on leur demande s'ils ont eu recours à ces pratiques.

de comités d'éthique ce que Poincaré (cité par Fagot-Largeault, 2010) aurait considéré comme une intrusion scandaleuse. Comme dans d'autres domaines, l'exacerbation de la compétition/concurrence impose le couple déviance/régulation.

Pour revenir à des fonctionnements plus courants du travail scientifique, la pression à la publication, à nouveau sans que l'on puisse en faire la preuve de manière mécanique, joue sur la taille des échantillons interrogés dans les enquêtes. Marszalek (2013) montre, qu'en dépit de la constitution d'un groupe d'experts en statistiques au sein de l'*American Psychological Association* et de leurs recommandations sur l'augmentation de la taille des échantillons pour assurer la fiabilité des inférences statistiques, le nombre moyen de sujets recrutés en psychologie expérimentale a diminué par rapport aux années 70 (41 en moyenne à cette époque contre 26 en 2006) en dépit des gains considérables de temps obtenus grâce à l'informatisation des passations des expériences et du traitement statistique.

3.2 Le traitement des données : du « data cooking » à l'auto-censure

Fanelli et Ioannidis (2013) montrent des liens plus explicites entre la pression à la publication et la tendance à la confirmation des hypothèses pour les données comportementales⁵ de recherches médicales. Les chercheurs d'universités américaines obtiennent des effets plus importants en faveur de leurs hypothèses que ceux constatés dans le même type d'études, réalisées par des chercheurs exerçant dans des pays où la bibliométrie n'a pas le même lien mécanique avec les diverses formes de

récompenses (primes, évolution de carrière, etc.). Le même effet n'est pas constaté sur d'autres mesures biologiques de type prélèvements etc. qui sont plus exposées à la réfutation pour reprendre les termes des auteurs.

Enfin, certains résultats non significatifs, ou concluant à l'absence d'effet, ne se publient pas ou peu. Comme pour la fraude, il est impossible de quantifier ce qui n'est pas publié pour des critères statistiques, que ce soit dans le champ médical (Fanelli, 2010) et, plus encore, en psychologie expérimentale où environ 95 % des résultats conduisent au rejet de l'hypothèse nulle (Sterling *et al.*, 1995). Or une absence d'effet est en soi un résultat particulièrement important dans le champ biomédical, en dépit parfois de l'avis des laboratoires qui ont financé la recherche. Cette sous-publication de résultats négatifs peut être interprétée de deux manières : la première et la moins optimiste serait de la considérer comme un indicateur de fraude ou au mieux de pratiques de « *data cooking* », la seconde consisterait à expliquer cette absence par la propension des chercheurs à davantage douter des résultats négatifs que positifs. Ferguson (2012) interprète l'attrait pour la confirmation comme une conséquence paradoxale de l'aspiration de la psychologie à être reconnue comme une « *as a purely objective science* », qui produit des preuves. Or l'absence de réfutation conduit selon lui à transformer les débats en psychologie en « batailles de boules de neige idéologiques », stériles au plan scientifique.

3.3 Effets de modes et modes d'argumentation

La volonté/nécessité de publier conduit à un certain conformisme dans le traitement statistique des données. Ainsi, dans le champ de la psychologie, l'analyse de variance s'est progressivement imposée au détriment de tests non-paramétriques. Les tests paramétriques,

⁵ En recherche biomédicale, on oppose données biologiques issues de prélèvement ou d'imagerie médicale aux données comportementales issues de l'observation des comportements des patients.

plus sensibles aux « cas marginaux », conduisent souvent, semble-t-il, à leur éviction des échantillons (Zimmerman, 1994). Ce choix, plus mû par les habitudes et les nécessités de se conformer à des normes de publication, conduit peut-être davantage à la cuisine des données, évoquée précédemment, qu'à un usage raisonné des statistiques inférentielles.

Des normes implicites similaires en analyse de données conduisent à des effets d'autres natures et non moins importants sur les productions scientifiques. Ainsi, l'usage dans les publications anglophones des méthodes de régression ont conduit à un abandon progressif de techniques plus descriptives comme les analyses de correspondances dans le monde francophone. Cette évolution n'est pas anodine au plan du raisonnement sous-jacent à ces traitements. Les méthodes de régression s'inscrivent dans un schème expérimental conduisant à réfléchir « toutes choses égales par ailleurs » alors que les méthodes d'analyse de données mettent en œuvre une rhétorique plus descriptive qu'explicative. Comme le montre Nétumières (1997), ces méthodes ne conduisent pas aux mêmes constats sur les mêmes données et n'appellent pas les mêmes conclusions ou interprétations. Desrosières (1995) souligne également cette différence, en montrant que les méthodes d'analyse de données invitent à décrire des cooccurrences, à penser par classes ou groupes sociaux, alors que les méthodes de régression conduisent à des discours plus mécanistes reprenant le vocabulaire des variables. Le raisonnement causal est plus tourné vers l'action ou le conseil, statut auquel prétendent implicitement de nombreuses recherches en sciences humaines. L'importance de ce que les chercheurs pensent être des attentes sociales à propos de la recherche en éducation conduit ainsi plus facilement à

l'explication, voire à la prescription, qu'à la description. Raisonner dans les termes des variables conduit aussi à des réifications de certains indicateurs, qui peuvent ensuite être mal compris ou perçus hors du champ de recherche qui les a produits.

Si tous les faits scientifiques sont débarrassés par la publication des conditions de leur élaboration, dans le champ de l'éducation, la définition mathématique qui est mise derrière les termes « inégalité d'accès à... » est un enjeu fondamental pour les conclusions que l'on peut tirer des recherches. Ce vocabulaire apparemment univoque « des chances d'accès » masque des descriptions statistiques parfois très différentes. Combessie (2011) démontre ainsi clairement que l'adhésion à une orthodoxie statistique ou une autre peut repeindre selon ses termes « en rose » ou « en noir » les mêmes données. Au-delà des conclusions et interprétations, les choix du mode de traitement des données vont jusqu'à modeler les théories. L'usage récurrent de régressions conduit à perdre de vue que « toutes choses égales » par ailleurs n'est qu'une hypothèse statistique et pas la description de la répartition des variables dans la population étudiée. Passeron (1991) rappelle en effet que la quête d'une perfection du raisonnement expérimental fondant les analyses de régression conduit à des « absurdités sociologiques ». Le comparatisme tous azimuts, suscité par les enquêtes internationales en éducation, conduit parfois à actualiser le paradoxe de Simiand qui s'interrogeait sur la manière dont vivraient les rennes dans le désert et inversement comment les chameaux s'adapteraient au grand Nord (cité par Passeron, 1991). Toute méthode d'analyse, toute interprétation des données est discutable et discutée, encore faut-il que publier ne soit pas conditionné par l'atteinte de seuils de significativité ou par l'emploi de certaines méthodes

d'analyses spécifiques. Là encore, les stratégies de publication poussent à une certaine homogénéisation potentiellement délétère au plan scientifique.

4. Écrire : la recherche ne s'exprime pas dans l'article, elle s'y réalise

4.1 Standardisation des publications

Dans la veine de ce qui vient d'être décrit pour le traitement statistique des données, la publication est conditionnée par la conformité à des « standards » plus ou moins contraignants. La structure IMRED⁶, issue des comptes rendus expérimentaux, s'est ainsi imposée en sciences biomédicales dans les années 1950 et plus tardivement en sciences humaines et sociales. Le choix de ce modèle, qui tient aujourd'hui de la convention de communication, n'est pourtant pas sans présupposés épistémologiques. Comme le décrit très clairement Pontille (2007), ce format a des effets conjointement sur plusieurs plans, de la mise en forme, pour citer le plus visible, à l'argumentation. L'homogénéisation des écrits scientifiques influe aussi sur les pratiques de lecture, le format IMRED permettant des lectures fragmentées. Les choix de mise en page faits par certaines revues en ligne comme « PlosOne » laissent supposer l'existence de ces stratégies de lecture. Les figures sont, en effet, détachées de l'ordre sériel de lecture et présentées sous forme d'un diaporama entre le résumé et le texte. Ce choix éditorial, vraisemblablement pensé pour son caractère ergonomique, montre à quel point ce format de présentation d'articles est tourné plus vers l'exposé de faits que vers une logique argumentative et plus encore interprétative.

⁶ Introduction Matériel/Méthode Résultats Et Discussion, en anglais IMRA(nd)D

La standardisation des écrits scientifiques peut être aussi vue comme une des conséquences de la croissance considérable du nombre d'articles scientifiques depuis les années 1950. En dépit de l'efficacité de ces standards pour la « chaîne » éditoriale, des chercheurs aux éditeurs, en passant par les relecteurs, elle conduit à une désingularisation au sens de Karpik (2007) propre à toute production de masse. Nombre d'auteurs en déplorent les effets transformant, selon eux, l'écriture en remplissage de formulaires (Pontille, 2007). Roland (2007, p.425) va plus loin en affirmant que le « *ready-to-write* » conduit au « *ready-to-think* ». La standardisation peut conduire à d'autres écueils : « à l'instar de tout dispositif standardisé qui sédimente des formes d'action dans des objets, il [le format IMRED] rend simultanément possible la copie et la contrefaçon. » (Pontille 2007, p. 14). Ce tableau des effets du formatage des publications pourrait passer pour un réquisitoire sans fondement, si la proposition de Pontille qui vient d'être évoquée ne se trouvait pas depuis confirmée⁷.

4.2 La standardisation de l'écriture

La standardisation des formats de publication s'accompagne également

⁷ Labbé (2013) a créé un outil de détection des articles écrits par un générateur automatique de faux articles (SciGen) en s'appuyant sur des outils de traitement automatique des langues et sur un corpus de faux articles élaborés pour comprendre les ressorts du calcul du h-index dans la base Google Scholar (Labbé, 2012). Sur près de 15000 publications téléchargées depuis des bases d'articles en informatique, 85 articles ont été identifiés comme étant produits automatiquement dont 16 dans la base IEEE commercialisée par Springer. En revanche, aucun article de ce type n'a été détecté dans l'archive ouverte arXiv. Ces articles n'auraient jamais dû passer la barre d'une relecture par les pairs, dans les 24 conférences victimes/complices de ces faux qui affichent des taux de sélection compris entre 18 et 30 %, ce qui montre la fragilité de ce type d'indicateurs comme signe de qualité.

d'évolutions dans l'écriture, dans ce que l'on pourrait qualifier de style si cette notion ne semblait pas hors sujet à propos de l'écriture scientifique. Pontille (2007, p.10) résume très clairement cet effet ainsi :

« Le récit expérimental convoite, non la littérarité de la lettre porteuse de l'empreinte de la personne, mais l'idéalité du sens. La mise en forme écrite est consacrée à la fabrication de « faits » par un travail d'articulation à d'autres contextes énonciatifs pour que s'opère l'effacement des conditions d'énonciation premières. Dans ce travail de représentation, le format IMRAD permet aux auteurs d'inscrire leur raisonnement dans le corps du texte tout en garantissant leur disparition nécessaire. ».

L'effacement énonciatif qui est une des caractéristiques distinctives du discours scientifique se traduit de deux manières : dans les marques d'énonciation et dans les manières de citer. Paradoxalement, alors que les scientifiques sont en quête permanente de visibilité, l'évolution des écrits concourt à une forme d'effacement de l'auteur. Pontille (2004) affirme que la présence de l'auteur dans le texte tend à être confinée dans la signature, place encore réduite par le nombre croissant de signatures.

Le caractère impersonnel des textes ne tient pas qu'à des formes passives ou à l'absence de marques linguistiques explicites ni encore à des emplois lexicaux du type *may/might* plutôt que *will*, il se traduit également par ce que Hyland (1998, cité par Roland 2007) désigne par « *hedges* ». Il s'agit de procédés rhétoriques qui permettent aux auteurs et lecteurs de négocier la légitimité des propositions notamment en mettant en jeu le renvoi à des références. Hyland précise que cette technique argumentative peut être

interprétée comme orientée vers la précision ou orientée vers l'auteur, pour lui éviter de trop s'engager. Certaines formulations ayant les dehors d'une expression précise ont en définitive pour but d'éviter tout conflit direct avec l'auteur : « In our study, the existence of xxx isoforms questions the validity of earlier results... ». Ces formulations courantes deviennent des tics d'expression d'autant plus facilement que, comme le souligne Roland (2007), la formation à l'écrit scientifique est réduite à une pédagogie du mimétisme. Ces précautions évitent des révisions trop nombreuses et par conséquent accélèrent la publication. Le dévoilement des « hedges » de *précisions* vers l'emploi de *précaution* (pour ne pas s'engager) doit être interprété selon Hyland (1998) dans un contexte de pression à la publication qui habille la confrontation de points de vue en confrontation de faits. L'emploi de nombreuses références, la justification de la moindre affirmation, même la plus générale, s'inscrit également dans cette logique de flottement entre une *recherche de précision* et un *retrait protecteur*.

Les raisons de citer sont multiples et ne relèvent pas exclusivement des stratégies d'écriture. En revanche, la manière de citer peut traduire des formes variées d'engagement dans le « dialogue » avec les discours rapportés. L'engagement le plus fort consiste à citer l'auteur au sens de reproduire un extrait plus ou moins long d'un texte [type 1]. Évoquer une expression, une notion, un concept, un fait employé ou relaté par un auteur est déjà plus distant mais toujours une désignation explicite [type 2]. Citer de manière allusive peut être produit principalement de deux façons, soit par l'évocation d'une notion ou d'un résultat suivie d'une parenthèse (avec un ou plusieurs noms) [type 3], soit par l'accolement de citations à une idée ne faisant pas référence au contenu précis du travail cité, mais par l'évocation de son

thème général.

“As observed at the turn of the century by Marks & Spencer (1899) [Type 2], who first named the “yelling reaction” (YR) [...]. Although numerous behavioral (Zeeg & Puss, 1931; Roux & Combaluzier, 1932; Sinon et al., 1948) [Type 3], pathological (Hun & Deu, 1960), comparative (Karybb & Szyla, 1973) and follow-up (Else & Vire, 1974) studies have permitted”⁸

Reste un cas particulier dans cette typologie par degrés d'engagement, celui d'expressions comme : « les saussuriens », le concept « foucaldien » de..., dans une perspective piagétienne... ». Les auteurs ne sont évoqués qu'à un niveau très général et pourtant la désignation de familles de travaux ou de courants de recherche représente une forme d'engagement forte pour l'auteur car il pourrait se disqualifier en employant mal à propos ces termes. Ce mode de citation, plus fréquent chez les auteurs chevronnés, pourrait passer même pour de la prétention pour un doctorant. Ces jugements témoignent bien de l'engagement que cette manière de convoquer d'autres auteurs suppose.

Enfin, pour clore sur cet aperçu des effets de la bibliométrie sur les pratiques d'écriture, il faut rappeler que la standardisation des bibliographies est une condition presque nécessaire à la bibliométrie. Le formatage des bibliographies est assez tardif dans les revues en sciences humaines francophones qui pratiquaient plus fréquemment la note en bas de page que la liste de référence finale.

4.3 Ce que citer veut dire : la double face des citations

Citer c'est juger et être jugé, à la manière dont Bourdieu (1982) considérait que parler c'était non seulement communiquer mais également s'assurer de certains profits sur le marché des échanges langagiers : être jugé positivement, être écouté, etc. Citer est fondamentalement un acte d'écriture et d'argumentation, mais c'est aussi un acte social d'inscription dans un réseau de recherches et d'idées et aussi de position dans le champ. À ce titre, il fait l'objet de nombreuses sollicitations plus ou moins coercitives des relecteurs, voire des comités de rédaction (Wilhite, 2012).

L'omniprésence de la bibliométrie comme signe de la qualité suscite parfois des réactions stratégiques caricaturales. Certaines revues ont en effet été exclues des bases de données pour avoir dopé artificiellement leur facteur d'impact en organisant des « cliques de citations » (Pontille, 2014). Ces comportements suscitent en retour la construction d'outils de détection de ce type de comportements, l'intervention de comités d'éthique, etc. Les dispositifs de quantification suscitent donc conjointement des comportements déviants et la nécessité d'une nouvelle intervention hétéronome, qualifiée parfois de bureaucratique par les chercheurs. Comme sur les marchés financiers, l'équilibre autorégulateur ne fonctionne que si tous les acteurs « jouent le jeu », dans les périodes de stabilité monétaire où la confiance dans la monnaie est importante (Orléan 2009). Les mêmes mécanismes semblent s'instaurer à propos de la valeur d'un article ou d'une citation : tout va bien tant que les auteurs ne s'écartent pas des normes de production des connaissances scientifiques.

⁸ Extrait de Perec, G. (2007). *Cantatrix sopranica L : Et autres écrits scientifiques*. Paris: Points. Parodie savoureuse d'un article IMRED

Les récents canulars ou opérations de testing de comités de lecture⁹ de revues jettent cependant le trouble sur la confiance dans l'évaluation par les pairs, et partant, sur la valeur des publications et des citations. A travers la fabrication d'un faux chercheur et de faux articles, Labbé (2010) a également montré les limites de ces indicateurs et surtout l'existence de faux articles. Cependant ce travail, qui propose une critique radicale de la bibliométrie, renforce paradoxalement l'idée d'une nécessaire régulation non pas de l'usage de la bibliométrie mais des pratiques de recherche et de publication. Ce type de critiques pour probantes qu'elles soient, finit par remettre en cause les pratiques de recherche plutôt que la bibliométrie. C'est du reste probablement le signe que la bibliométrie est considérée comme externe à ces affaires de fraudes et que la responsabilité ne peut en être imputée qu'aux chercheurs « déviants » ou stratèges.

Sans en sous-estimer ses effets, Gingras (2014) défend que l'augmentation continue du nombre de références par articles correspond à l'augmentation de la population des chercheurs et que le nombre d'autocitations n'a pas beaucoup augmenté depuis les années 1980. Ce n'est donc pas tant l'effet au sens statistique qui est inquiétant que l'apparition de ces phénomènes stratégiques qui remettent en cause les modes de régulation du champ scientifique.

L'allongement des listes bibliographiques est, du reste, peut-être moins un effet direct de la bibliométrie évaluative que de la diffusion via des modèles comme l'IMRED des pratiques d'écriture, issus des recherches biomédicales. Cependant, multiplier les références dans une citation

allusive, dans un compte-rendu d'essai clinique n'a pas le même sens qu'en sciences humaines et sociales, de même que compter le nombre d'études attestant l'effet d'un traitement n'a pas le même sens que d'évoquer les résultats convergents d'expériences en psychologie. L'homogénéisation des modes d'écriture conduit à des transpositions hasardeuses de pratiques d'écriture perdant leur signification d'origine.

Toujours dans le but de dépasser les a priori sur les pratiques de citation, Gingras (2014) montre clairement qu'il n'y a pas de génies méconnus, que la plupart des chercheurs reconnus sont abondamment cités. Inversement, il est intéressant de constater que les chercheurs reconnus ne sont pas exempts de stratégies bibliographiques que leur statut devrait leur éviter. Le titre d'un article de de Singly (1998) résume habilement ces stratégies : « Bourdieu : nom propre d'une entreprise collective. ». Enfin si la bibliométrie semble éviter les erreurs de type II (chercheur reconnu mais peu cité), elle ne semble pas à l'abri de l'autre type d'erreurs qui consisterait à accorder un crédit bibliométrique à des travaux d'une qualité plus discutable.

Au-delà des constats globaux produits par la bibliométrie « sociologique » et de leurs interprétations, la multiplicité des pratiques maintes fois attestées, telles que les publications « salami », ou encore les dénonciations de chercheurs invités à signer alors qu'ils ne sont pas auteurs, etc. attestent pourtant bien de l'existence d'une préoccupation, voire d'une obsession, pour cette visibilité quantitative. L'augmentation très importante, quelle que soit la discipline (Gingras, 2014), peut également faire l'objet d'une interprétation ambivalente : est-ce le fruit d'une « collectivisation » de la recherche ou la somme d'intérêts individuels ? Ce constat a conduit, par

⁹ Voir le numéro spécial de la revue Science d'octobre 2013 : « *Communication in Science: Pressures and Predators* »

exemple, à une révision de la valeur d'une publication en fonction du nombre d'auteurs ; en imposant une règle pour tous, cette réaction est symptomatique des tâtonnements de la bibliométrie évaluative et de l'absence d'analyse des spécificités disciplinaires. Les pratiques stratégiques d'optimisation liées à la bibliométrie qui reposent sur le fait de prendre des indicateurs pour des objectifs ne sont profitables qu'à très court terme, car l'inflation du nombre de publications et des citations conduira nécessairement à une réévaluation des seuils pour être recruté par exemple, donc à une forme de dévaluation de la monnaie-citation. Ce mécanisme n'est pas propre au champ scientifique, et bien qu'il ait été maintes fois attesté (Beauvallet 2009) et dénoncé pour ses effets délétères, il ne freine pas la quantophrénie !

4.4 La monnaie d'academicus

Citer n'est plus seulement un acte de rhétorique scientifique mais est aussi devenu un acte bibliométrique. La citation garde, de son côté face, les caractéristiques d'un bien symbolique échangé dans une économie du don contre-don de la reconnaissance scientifique, mais, de son côté pile, elle est l'unité fondamentale de la bibliométrie. Les citations sont devenues une forme de monnaie et assurent dans le champ scientifique les trois fonctions dévolues aux monnaies dans une conception classique (Blanc, 2009) : fonction de compte (unité de mesure de la qualité par l'audience), d'intermédiaire des échanges (valorisation des publications dans les demandes de fonds ou les recrutements, voire indexation du salaire sur l'audience des publications) et de réserve (accumulation d'un capital symbolique qui se traduit notamment par l'augmentation avec l'âge d'un indicateur comme le h-index).

Les citations, comme la monnaie, ne peuvent être crédibles sans confiance,

comme le souligne le terme fiduciaire. L'assentiment général de reconnaissance de la citation comme indicateur de la qualité scientifique est le fondement de la confiance dans cette monnaie. Certaines pratiques de publication ont tendance à dévaloriser la monnaie ; s'instaurent alors des formes de défiance qui conduisent à produire des indicateurs bibliométriques comparables, dans une certaine mesure, au rôle d'assurance de la valeur qu'on fait jouer à des indices économiques comme l'indice des prix ou des taux de change (Orléan, 2009). Les éditeurs scientifiques ont joué le rôle d'instance de régulation des marchés lorsque des revues se sont organisées en cliques pour se citer mutuellement.

Enfin comme Orléan (2009) le décrit, la monnaie est désirée en raison de son pouvoir d'achat, tout comme la citation pour ce qu'elle apporte la renommée scientifique. Le prix est bien, sur certains marchés où la détermination de la valeur est incertaine, un signe de qualité. Tant de similitudes conduisent à supposer que la bibliométrie transforme une économie du don, qu'est le jeu d'hommages réciproques véhiculé par les citations, en une économie de marché.

5. Conclusion : La surface des faits et les effets plus profonds

5.1 Du comparable pour comparer : effets d'homogénéisation

L'enjeu principal de ce travail était de décrire des évolutions corrélatives à l'introduction de la bibliométrie évaluative. Au-delà des effets caricaturaux qui témoignent explicitement d'une influence de ces dispositifs de quantification, l'évocation de choix plus discrets, tous orientés implicitement vers l'intention/nécessité de publier, avait pour but de provoquer une distanciation sur les pratiques d'écriture qui participent au développement et à la légitimation de

la bibliométrie évaluative. Le propos n'est pas de défendre l'existence d'effets cachés, ourdis au profit d'une domination culturelle ou d'éditeurs scientifiques, mais de rappeler que les « allant-de-soi » en termes de normes de productions scientifiques sont des conventions socialement et historiquement construites. Les chercheurs participent, par leurs choix, à la légitimation de la détermination de la valeur scientifique par l'audience sous-jacente à la bibliométrie évaluative, ce qu'ils perçoivent et décrivent paradoxalement comme une intrusion.

La bibliométrie repose nécessairement sur une forme d'homogénéisation des productions, évoquée en introduction, pour laquelle il reste bien difficile d'affirmer de manière globale qu'elle en est la condition ou la conséquence. Pour tous les éléments des écrits scientifiques indexés dans les bases de données, la normalisation est une condition de la bibliométrie. En revanche, l'homogénéisation du format de présentation (plan, organisation de l'article) n'est de ce fait pas une condition de la bibliométrie mais l'est en revanche de la publication. En effet, certaines revues imposent des formats de soumission et la mise en ligne des revues impose également des formes d'homogénéisation des présentations en dépit de la liberté qu'offre l'édition électronique.

Comme l'a décrit Bourdieu (1996) à propos de la course à l'audience dans les médias, la bibliométrie évaluative, qui est essentiellement une mesure d'impact, pousse à la quête de la primauté qui n'est pas nécessairement de la nouveauté et moins encore de l'originalité. « Pour être le premier à voir et à faire voir quelque chose, on est prêt à peu près à n'importe quoi, et, comme on se copie mutuellement en vue de devancer les autres, de faire avant les autres ou de faire

autrement que les autres, on finit par faire tous la même chose, la recherche de l'exclusivité, qui, ailleurs, dans d'autres champs, produit l'originalité, la singularité, aboutit ici à l'uniformisation à la banalisation » (Bourdieu, 1996, p. 20). Présentée hors de son contexte (à propos du journalisme télévisuel), cette description pourrait passer pour un extrait d'un des nombreux réquisitoires contre la bibliométrie évaluative.

Toujours dans la veine des effets d'audience, la bibliométrie contrairement à son apparence de marché ouvert où tout auteur peut émerger, accroît l'effet Matthieu déjà décrit par Merton en 1968. Ces mouvements d'accumulation, déjà connus de longue date des bibliométriciens pour les citations, sont amplifiés par l'omniprésence, y compris dans les outils de documentation, des statistiques de citations et autres indicateurs d'impact. Le lien mécanique parfois établi par différentes institutions d'évaluation et de financement de la recherche contribue à convertir l'audience en crédits de recherches et par conséquent à favoriser l'expansion des recherches d'un laboratoire donné qui seront plus citées, etc. Les effets Matthieu contribuent à définir des champs de recherche légitimes, novateurs qui susciteront l'intérêt d'autres chercheurs et contribueront encore à citer plus les travaux des « pionniers » déjà beaucoup dotés.

La preuve la plus manifeste de l'homogénéisation due à la bibliométrie est que tout vaut un. Si la monnaie bibliométrique est la citation, la simple évocation (citation de type³) d'une recherche valorise autant un article (par une citation) qu'une critique approfondie et documentée. Cette réduction qui au départ était une invention documentaire¹⁰

¹⁰ Le SCI a été conçu sur le modèle du *Shepard's Citor* qui recensait la jurisprudence aux Etats-Unis (Wouters 2006)

devient une redéfinition de la valeur¹¹. Aux éternels débats épistémologiques sur les conditions de validité des recherches, le pragmatisme bibliométrique substitue une définition de la valeur par une tautologie : si le nombre de citations mesure la qualité, « *the most highly cited scientists are highly cited because they are highly cited (= good as indicated by the high citation counts)* » (Luukkonen, 1990, cité par Wouters 1999).

5.2 Ethiques scientifiques et marchandes : une redéfinition de la valeur.

La définition de la valeur de manière autoréférentielle est caractéristique du fonctionnement de marchés financiers en crise, où, lorsque personne ne sait ce que vaut un titre, on se fie à l'opinion commune. Orléan (2009) décrit ce mode de fonctionnement en référence à la métaphore du concours de beauté proposée par Keynes. Certains journaux proposent aux lecteurs de dire quelles seront les six photos les plus choisies par l'ensemble des lecteurs. Il n'est alors plus question d'évaluer « substantiellement » mais d'anticiper les évaluations des autres. La valeur est donc définie par ce que Karpik (2007) qualifierait de l'opinion commune. Ce mode de jugement qui ne repose pas sur une évaluation de la valeur scientifique à partir du contenu, rend possible une caractéristique importante de la bibliométrie évaluative : elle permet à ceux qui ignorent tout d'un domaine de recherche de déterminer qui en sont les acteurs importants. C'est du reste en partie ce qui minimise sa portée aux yeux des chercheurs, car, en consultant les statistiques d'impact des revues ou des auteurs de leur domaine, ils n'apprennent rien et à ce titre ces données leur

paraissent inoffensives.

La définition de la qualité par l'audience est une intrusion forte dans le champ scientifique qui écrase les principes laborieusement décrits et construits par l'épistémologie des conditions de validité toujours incertaine d'une recherche. Cette redéfinition de la valeur est la marque d'une forme de transition d'un ordre aristocratico-institutionnel vers un ordre marchand. Cette évolution s'accompagne d'une « mise en examen » des pratiques et normes fondatrices pour l'autonomie et la validité des travaux de recherche. Opposer l'éthos scientifique décrit par Merton à une « éthique » marchande est toutefois simplificateur ou caricatural pour deux raisons. D'une part, le champ scientifique ne s'organise pas strictement comme le décrivent les normes mertonniennes, pas plus que le marché ne fonctionne conformément à l'équilibre général décrit dans l'économie néo-classique. D'autre part, la recherche n'a pas basculé complètement ou brutalement dans les principes de l'éthique marchande.

Cependant, cette confrontation permet de mettre en avant des évolutions qui accompagnent la généralisation de l'usage évaluatif de la bibliométrie. Ainsi, en référence à l'éthique marchande, on se méfie des cooptations (fonctionnement pourtant courant et essentiel dans le champ scientifique) qui passent pour de l'entente ou de la subordination (sont deux formes d'entrave à la libre concurrence). De la même manière, les diverses formes de collaborations, de co-signatures pourraient passer pour conduire à une forme de paresse sociale, pour amener une suspicion (pour une part légitime) sur la participation effective des auteurs aux recherches. L'éthos marchand conduit plutôt à exacerber la compétition individuelle dans l'espoir de faire baisser les prix.

La communication scientifique repose sur

¹¹ À titre d'exemple, on trouve ce type d'affirmation sur le site harzing.com : « *The advantage of the h-index is that it combines an assessment of both quantity (number of papers) and quality (impact, or citations to these papers)* » (Glänzel, 2006).

une confiance a priori sur la qualité du recueil des données et la sincérité des analyses. Cette confiance, notamment garantie par les institutions, est à l'opposé de ce que les comptables appellent le *window-dressing*. Cette pratique qui consiste à arranger les comptes en masquant ou valorisant certaines plus-values pour favoriser les cours de bourses, suscite au contraire une méfiance a priori dans les données et les statistiques. Certains courants de la comptabilité défendent même que « c'est le cours de bourse qui mesure le mieux la réalité économique de la firme, car les marchés sont réputés efficaces » (Chiapello et Desrosières, 2006, p. 303), ce qui représente un rapport pour le moins singulier à la quantification en tant que représentation du réel. Transposés à la recherche, ces principes conduiraient à juger de la valeur scientifique en suivant le cours des citations plutôt qu'en évaluant la validité et la portée des résultats. L'incertitude sur la valeur conduit, dans une logique marchande, à créer des instances de certification, d'évaluation. Faute de pouvoir estimer de manière fiable la valeur d'un actif, on multiplie les labels, les normes censées en garantir la qualité. Dans cette logique de détermination de la valeur, diverses agences, outils d'évaluation se multiplient, des listes de revues accordent des formes de labels. L'usage évaluatif des facteurs d'impact est typiquement de cette nature, en privilégiant la valeur des supports de publications plutôt que de juger de la qualité d'une recherche. Les évolutions récentes en France relèvent de cette culture marchande : individualisation des récompenses pour favoriser une compétition prétendument stimulante, excellence définie par la rareté, suspicion d'ententes sur le recrutement des enseignants qui a imposé comme garantie évidente d'objectivité, l'intervention des membres externes à la discipline dans les comités de recrutement, etc.

5.3 L'impuissance de la critique : « one number beats no number¹² »

En dépit des très nombreuses critiques émises sur les usages et les pratiques de la bibliométrie, elle s'impose dans les outils de documentation, dans les comités de recrutement ou les instances attribuant les promotions : pour ce qui compte, il faut compter ! Le peu de prise de la critique peut et doit être imputé à l'obsession actuelle pour la quantification et au poids des chiffres face à tout argument. N'importe quel chiffre vaut mieux que pas de chiffre du tout quand la concurrence généralisée passe pour le principal ressort de motivation. Des processus psycho-sociaux complexes motivent probablement les chercheurs à surveiller par des alertes les nouvelles citations de leurs articles tout en méprisant le « réductionnisme bibliométrique ».

L'adhésion plus ou moins volontaire des chercheurs à la bibliométrie évaluative tient paradoxalement aussi à des similitudes entre l'éthique universitaire et marchande, qui pourraient leur laisser croire à une forme de compatibilité voire de légitimité des valeurs qu'elles véhiculent. Si la manière dont est établie la valeur et la confiance dans le monde scientifique est a priori radicalement opposée à ce qui a été décrit du fonctionnement des marchés, les chercheurs sont sensibles aux valeurs démocratiques véhiculées par le marché. L'insubordination¹³ propre au marché rejoint clairement l'autonomie lentement conquise et tant chérie des universitaires. Ce qu'Orléan (2009) qualifie de séparation marchande (conséquence de l'insubordination) conduit, selon lui, les acteurs à une quête permanente de

¹² Gingras (2014) résume par cette formule éloquent le fétichisme numérique.

¹³ Aucun vendeur ne peut imposer d'acheter et inversement.

visibilité. La volonté de susciter l'intérêt d'autrui est également une préoccupation fondamentale des chercheurs qui voient tous dans leur sujet ou leur mode d'explication la clé de différents problèmes.

L'idée de marché véhicule aussi une forme de démocratisation de l'accès au système de récompense. Matalon et Lemaine (1960), décrivaient justement qu'en science, il n'y pas de second. La bibliométrie semble réformer cet élitisme des grands prix scientifiques. Elle instaure non-seulement un second, mais tout une métrique permettant des comparaisons y compris transdisciplinaires, ce qui était inaccessible antérieurement. L'élitisme est à portée de clics : on peut télécharger son certificat¹⁴ de l'article le plus cité dans le top 25 d'Elsevier ! À l'insubordination s'ajoute l'égalité supposée des acheteurs-vendeurs, qui correspond assez bien au rôle biface d'auteur-relecteur qu'ont les chercheurs. Cette égalité de statut présumée minimise aussi les effets de stratifications sociales, laissant supposer que « tout le monde a sa chance » ou du moins doit jouer le jeu de la bibliométrie.

Cette ambivalence, face aux valeurs marchandes, est peut-être une des pistes pour comprendre la faible portée des critiques de la bibliométrie évaluative et la poursuite de la course effrénée à la publication qu'elle suscite. Enfin, la massification de l'enseignement supérieur n'est peut-être plus compatible avec l'élitisme aux valeurs aristocratiques qui a structuré l'émergence des sciences depuis la fin du XVIIIe siècle. Les modes de reconnaissance et de récompense ne sont peut-être plus adaptés à ces nouvelles échelles. Il reste à inventer des modes de régulation qui concilient les valeurs fondamentales, qui rendent possible la production de connaissance (jugement

« sur pièce », autonomie, etc.), et les conditions toujours changeantes de l'exercice du travail scientifique.

5.4 Lire, écrire, commenter ou compter, liker, classer : quelques tendances et propositions pour faire des choix

S'il est probablement difficile de lutter contre ce mouvement vers un ordre marchand dont l'extension dépasse le champ scientifique, un enjeu pourrait être de déterminer les conditions permettant de restaurer la science comme conversation vive et loyale. La numérisation des supports de publication pourrait être l'occasion de réfléchir sur les choix de publication mais aussi des usages de la bibliométrie. Puisque tout ce qui ne se compte pas ne vaut rien, il faut compter ce qui est caché et pourtant fondamental dans le système « classique de publication »

- Des expériences de publications partielles de rapports des relecteurs permettent de valoriser l'évaluation par les pairs qui est à la fois à la base de la conversation scientifique mais aussi une condition essentielle à l'existence de la bibliométrie via l'accréditation des publications. Si, par l'inflation des publications et l'effacement de ce travail important, on est conduit à demander aux relecteurs de cocher des grilles, on ne pourra qu'aboutir à remettre en cause leur jugement notamment à l'aune de leur divergences comme ce fut le cas dans l'évaluation scolaire. Or la supériorité de l'évaluation par les pairs est précisément d'enrichir les publications, d'être subjective mais fondée, personnelle mais au nom d'une communauté scientifique. Le rôle des relecteurs est fondamental et insuffisamment valorisé. Ils sont partiellement co-auteurs des écrits scientifiques et parfois plus que

¹⁴

<http://webshop.elsevier.com/certificates/Top25/>
consulté le 4 nov. 2014

certains directeurs de recherche ou de laboratoire qui signent les publications. Par ailleurs, leur activité structure les pratiques de publication dans leur champ de recherche, d'où le danger que représente l'appauvrissement de la relecture par le flot d'articles soumis qui a parfois pour conséquence l'utilisation de procédés de cotation sur des grilles plus rapides que des relectures approfondies.

- Dans un autre registre, la numérisation des revues pourrait être l'occasion de promouvoir la publication de données brutes associées à des articles. Le coût important de la collecte devrait inciter à ne pas en faire un trésor personnel, mais au contraire une ressource commune au même titre que les notions ou concepts. La production d'analyses sur des données brutes publiées serait un nouveau type de publication, qui conduirait peut-être à d'autres rapports aux données que le « *data cooking* » allant du *window-dressing* à la fraude.
 - Les publications en ligne n'étant pas régies par les mêmes contraintes, elles pourraient bousculer les standards établis en ne limitant pas le nombre de pages pour éviter les publications dites « salami », ou encore en ouvrant à la publication de médias associés. Le nombre considérable de conférences mises en ligne témoigne de cette évolution des modes de communication scientifique à grande échelle.
 - Les signatures différenciées des articles (écriture, mise au point du matériel, traitement des données...) sont probablement un moyen efficace d'éviter bien des brouilles, responsabiliseraient précisément les auteurs d'une partie de la publication et clarifieraient les spécialités ou domaines de compétences de chacun.
- Les grandes découvertes, particulièrement dans certains domaines où l'instrumentation est fondamentale comme en physique de la matière, ne seront que collectives. La signature de collectifs qui s'impose dans ce champ de recherche est probablement un puissant moyen de ramener la valorisation individuelle vers la participation à la recherche plus que vers la seule publication. Faire partie d'un collectif serait plus valorisé que de compter le nombre de publications ou de citations.
- Le travail de lecture pourrait être valorisé à travers d'autres usages et pratiques de la bibliographie. On pourrait inventer des moyens de différencier les citations selon leur statut dans le texte et mettre en références des recherches qui ne sont pas citées mais qui ont fait partie du travail de recherche : quelques indications ajoutées par les auteurs en hyperlien sur les références en bibliographie. Ce travail de qualification pourrait nourrir la recherche documentaire d'autres chercheurs de manière plus précise que les algorithmes de recommandations proposés par certains portails documentaires qui repèrent les cooccurrences, et conduisent nécessairement à un enfermement ou un renforcement des mises en relation. En cliquant sur une recommandation, on lui accorde un poids nouveau dans « ses » chances d'être recommandée ultérieurement.
 - L'édition en ligne pourrait être l'occasion de moins répéter les mêmes discours via la publication de liens permettant de suivre, en quelques clics, un auteur, une recherche d'article en article. Des tendances de ce type émergent dans les dépôts d'archives, ce qui rejoint certains projets de mise

en réseau d'articles sur le principe des épi-revues.

Les profonds et très récents bouleversements dans l'édition scientifique, comme l'accroissement du nombre de revues en accès libre, sont l'occasion de repenser les modèles de publication. Les dispositifs techno-éditoriaux qui viennent d'être décrits, multiplient les occasions de dialogue et donc nécessairement la production de connaissances nouvelles. Cependant, la numérisation peut tout aussi bien promouvoir l'audience (supposée) et son cortège de mesures comme finalité et valeur de l'activité scientifique. Est-il raisonnable de continuer à monologuer (puisque la plupart des articles sont peu lus), de manière stéréotypée (du traitement de données à l'écriture), de manière intensive (à un rythme appauvrissant toutes les lectures et relectures essentielles à l'activité de recherche), sur des sujets en vogue (ou produisant des mots clés attractifs) en quête d'une audience en partie construite d'autocitations et de citations de connivences ?

6. Références

- Beauvallet, M. (2009). *Les stratégies absurdes : Comment faire pire en croyant faire mieux*. Paris: Seuil.
- Blanc, J. (2009). Usages de l'argent et pratiques monétaires. In François Vatin et Philippe Steiner (Eds.), *Traité de sociologie économique* (pp. 209-246). Paris : Presses Universitaires de France.
- Bourdieu, P. (1982). *Ce que parler veut dire : L'économie des échanges linguistiques*. Paris : Fayard.
- Bourdieu, P. (1996). *Sur la télévision*. Paris : Raisons d'agir.
- Bourdieu, P. (2001). *Science de la science et réflexivité*. Paris: Raisons d'agir.
- Callahan, M, Wears, RL, & Weber, E. (2002). Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. *JAMA*, 287(21), 2847-2850.
- Carbone, P. (2010). *Coûts, bénéfices et contraintes de la mutualisation des ressources électroniques : éléments de comparaison internationale et propositions* (No. 2010-012). Paris : Inspection générale des bibliothèques.
- Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. *Dual-process theories in social psychology*, 73-96.
- Chiapello, È., & Desrosières, A. (2007). La quantification de l'économie et la recherche en sciences sociales : paradoxes, contradictions et omissions. Le cas exemplaire de la positive accounting theory. *Recherches*, 297-310.
- Combessie, J.-C. (2011). Analyse critique d'une histoire des traitements statistiques des inégalités de destin. *Actes de la recherche en sciences sociales*, 188(3), 4-31.
- Desrosières, A. (1995). Classer et mesurer : les deux faces de l'argument statistique. *Réseaux*, 13(71), 11-29.
- Dominguez, M., Vieira, M.-J., & Vidal, J. (2012). The Impact of the Programme for International Student Assessment on Academic Journals. *Assessment in Education: Principles, Policy & Practice*, 19(4), 393-409.
- Fagot-Largeault, A. (2013). Petites et grandes fraudes scientifiques : Le poids de la compétition. In F. Gérard (Éd.). *La mondialisation de la recherche : Compétition, coopérations, restructurations*. Paris : Collège de France.
- Fanelli, D. (2009). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PLoS ONE*, 4(5), e5738.
- Fanelli, D. (2010). « Positive » Results Increase Down the Hierarchy of the Sciences. *PLoS ONE*, 5(4), e10068.
- Fanelli, D. (2013). Why Growing Retractions Are (Mostly) a Good Sign. *PLoS Med*, 10(12), e1001563.
- Fanelli, D., & Ioannidis, J. P. A. (2013). US studies may overestimate effect sizes in softer research. *Proceedings of the National Academy of Sciences*, 110(37), 15031-15036.
- Ferguson, C. J., & Heene, M. (2012). A Vast Graveyard of Undead Theories Publication Bias and Psychological Science's Aversion to the Null. *Perspectives on Psychological Science*, 7(6), 555-561.
- Gingras, Y. (2014). *Les dérives de l'évaluation de la recherche : Du bon usage de la bibliométrie*. Paris : Liber.
- Gosselain, O. (2011). *Slow Science - La désexcellence*. Uzance, 1, 128 -140.

- Karpik, L. (2007). *L'économie des singularités*. Paris : Gallimard.
- Labbé, C. (2010). Ike Antkare one of the great stars in the scientific firmament. *International Society for Scientometrics and Informetrics Newsletter*, 6(2), 48–52.
- Labbé, C., & Labbé, D. (2013). Duplicate and fake publications in the scientific literature: how many SCiGen papers in computer science? *Scientometrics*, 94(1), 379-396.
- Marshall, E. (2000). Scientific misconduct. How prevalent is fraud? That's a million-dollar question. *Science*, 290(5497), 1662-1663.
- Marszalek, J. M., Barber, C., Kohlhart, J., & Holmes, C. B. (2011). Sample size in psychological research over the past 30 years. *Perceptual and Motor Skills*, 112(2), 331-348.
- Matalon, B., & Lemaine, G. (1969). La lutte pour la vie dans la cité scientifique. *Revue française de sociologie*, 10(2), 139-165.
- Medamar, P. B. (1964) Is the scientific paper a fraud ? In D. Edge (Ed.), *Experiment*. London: BBC Publications.
- Mehler, J., & Bever, T. G. (1967). Cognitive Capacity of Very Young Children. *Science*, 158(3797), 141-142.
- Nétumières, F. des. (1997). Méthodes de régression et analyse factorielle. *Histoire & Mesure*, 12(3), 271-297.
- Orléan, A. (2009). La sociologie économique de la monnaie. In François Vatin et Philippe Steiner, *Traité de sociologie économique* (pp209-245). Paris : Presses Universitaires de France.
- Passeron, J.-C. (2006). *Le Raisonnement sociologique: Un espace non poppérien de l'argumentation*. Paris : Albin Michel.
- Piaget, J. (1968). Quantification, Conservation, and Nativism. *Science*, 162(3857), 976-979.
- Pons, X. (2011) *L'évaluation des politiques éducatives*. Paris : Presses Universitaires de France.
- Pontille, D. (2007). Matérialité des écrits scientifiques et travail de frontières : le cas du format IMRAD. In P. H. et M. Paul-Cavallier (Éds.), *Sciences et frontières*, (pp. 229-253).
- Pontille, D. (2014). *Jeux de données et manufactures bibliométriques*. Séminaire : Big data, entreprises et sciences sociales - Usages et partages des données numériques de masse, Collège de France Paris.
- Roland, M.-C. (2007). Publish and perish. Hedging and fraud in scientific discourse. *EMBO Reports*, 8(5), 424-428.
- Singly, F. (1998). Bourdieu : nom propre d'une entreprise collective. *Le Magazine littéraire*, 369, 39-44.
- Sterling, T. D., Rosenbaum, W. L., & Weinkam, J. (1995). Publication Decisions Revisited: The Effect of the Outcome of Statistical Tests on the Decision to Publish and Vice Versa. *The American Statistician*, 49(1), 108.
- Wilhite, A. W., & Fong, E. A. (2012). Coercive Citation in Academic Publishing. *Science*, 335(6068), 542-543.
- Wouters, P. (1999). *The Citation Culture*. Ph.D. Thesis. Amsterdam: University of Amsterdam.
- Wouters, P. (2006). Aux origines de la scientométrie : La naissance du Science Citation Index. *Actes de la recherche en sciences sociales*, 164(4), 11.
- Zimmerman, D. W. (1994). A note on the influence of outliers on parametric and nonparametric tests. *Journal of General Psychology*, 121(4), 391.