



HAL
open science

Peut-on réenchanter le processus de publication ?

Héloïse Berkowitz

► **To cite this version:**

Héloïse Berkowitz. Peut-on réenchanter le processus de publication ?. Le Libellio d'AEGIS, 2016, 12 (2), pp.41-47. hal-03544105

HAL Id: hal-03544105

<https://hal.science/hal-03544105>

Submitted on 21 Feb 2022

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.

Peut-on réenchanter le processus de publication ?

Héloïse Berkowitz

i3-CRG École polytechnique CNRS Université Paris-Saclay

Le 5 avril 2016, le Centre national de la recherche scientifique (CNRS) a publié un avis intitulé *Discussion et contrôle des publications scientifiques à travers les réseaux sociaux et les médias : questionnements éthiques* portant, entre autre, sur les failles et conséquences négatives du système de *peer-reviewing* des articles scientifiques : explosion du volume d'articles soumis, accroissement du nombre de revues, multiplication des fraudes scientifiques et des rétractations d'articles, etc.

Étonnamment, l'avis du CNRS ne discute nulle part les conséquences d'une de ses propres initiatives, le classement des revues tel qu'il est par exemple mis en oeuvre dans la section 37 (Économie et gestion). Les effets pervers des classements et autres *rankings* d'écoles ou de journaux académiques sur les pratiques de publication des chercheurs ont pourtant largement été décrits et ont fait l'objet de nombreuses critiques (Osterloh & Frey, 2015 ; Kogut, 2008 ; Charroin, 2013 ; Osterloh *et al.*, 2008). C'est l'idée de technologie invisible développée par Michel Berry (1983). Dans cette perspective mécaniste, les individus maximisent ce sur quoi ils sont évalués et une fois que l'on met en place des systèmes d'évaluation ou de jugement, les acteurs vont s'y conformer mécaniquement (Riveline, 1991). S'il peut exister des apprentissages et des explorations possibles à partir des outils de gestion (Moisdon, 1997), les facteurs d'impacts et autres H index, mais aussi les primes à la publication, constituent autant d'instruments qui vont structurer les comportements dans le monde académique, depuis le recrutement jusqu'au choix des thèses de doctorat en passant par les appels à projets.

Les critiques ne s'arrêtent pas là puisqu'elles touchent aussi les maisons d'édition des revues académiques, accusées de gonfler leurs bénéfices sur le compte de la recherche, mais aussi de mettre les journaux et les relecteurs sous pression (aller toujours plus vite au détriment de la qualité et du travail de fond), ou d'encourager à la seule publication de résultats positifs, sans assurer la moindre gestion des risques (Springer a dû rétracter 64 articles en 2014).

Un certain esprit de désenchantement et d'aigreur souffle alors sur le monde de la recherche, et peut-être quelques pratiques alternatives de la publication permettront-elles de le réenchanter un peu.

Entre *dark* et *open*, les pratiques alternatives de la science collaborative

L'ouverture des données et l'accès gratuit à la connaissance constituent un enjeu contemporain majeur, mis en avant par les politiques et le monde académique, rarement par ceux qui pratiquent vraiment cette ouverture.

D'irréductibles Gaulois défendent des initiatives d'*open science*, dans lequel le système de relecture est ouvert et non anonymisé (*open peer review*). Y publier relève presque d'un acte engagé, politique, puisque les auteurs font dès lors une croix sur le classement de la revue, les classements de type CNRS ne prenant en compte que les revues à comité de lecture en double aveugle censées assurer plus de rigueur scientifique par l'anonymat (aucun biais ne serait introduit dans la relecture puisque, au moins officiellement, on ne sait qui est l'autre). Pourtant, il est aisé d'imaginer qu'au contraire, la visibilité permette d'éviter certains écueils notamment celui de la mauvaise critique d'un lecteur pressé ou agacé et forcerait au contraire à une relecture de fond, soignée et soigneuse, attentive à l'autre à cause d'effets réputationnels. Si aucun système n'est exempt de dérives possibles cela ne devrait pas empêcher d'expérimenter des solutions alternatives.

L'autre dimension de la science collaborative est la diffusion des articles par des voies différentes, qui contournent le système payant des éditeurs de revues. Certains forums mettent ainsi en contact des demandeurs d'articles (ceux qui n'ont pas les moyens d'y accéder par leurs institutions) et des offreurs d'articles qui leur répondent (ceux qui y ont accès par leurs institutions). Le modèle est poussé à l'extrême, voire au-delà des frontières du droit, avec le site Sci-Hub, la première plateforme pirate créée par la neuroscientifique russe Alexandra Elbakyan, qui met à disposition gratuitement des millions d'articles de recherche.

Refuser le système des maisons d'édition constitue également une alternative crédible dans certaines disciplines qui ont su résister à la financiarisation de la production des connaissances scientifiques.

Repenser la relecture : de la pré à la post publication

Le système de relecture en double aveugle est au cœur des débats et la cible des principales critiques, notamment parce qu'il n'empêche ni la fraude ni la publication d'articles creux (quand il ne l'encourage pas, Laroche, 2015). Différentes alternatives existent.

Avant la publication, le PPRI (Pre-publication Independent Replication)¹ est un système de test des résultats avant la publication des articles de recherche (Schooler, 2014), dérivé du problème de non répliquabilité des résultats dans bien des disciplines, majoritairement, en économie et en psychologie (Open Science Collaboration, 2015). L'idée, simple et de bon sens, consiste à vérifier la robustesse des résultats avant même qu'ils ne soient diffusés, dans des laboratoires indépendants – ce qui soulève des questions de coût et d'impartialité.

Le « Pipeline Project » est un projet de l'INSEAD qui a soumis à un test de répliquabilité dix hypothèses étudiées en laboratoire. Par exemple, a été sélectionné pour test, le « *Burn-in-hell effect* », selon lequel les participants pensent que des dirigeants sont plus susceptibles de brûler en enfer que des membres de catégories sociales définies par un comportement anti-social (comme les casseurs). Le test a été effectué dans vingt-cinq laboratoires indépendants. 60 % des hypothèses ont passé le test de répliquabilité. Deux ont échoué entièrement. Une a fonctionné aux États-Unis

1. <http://retractionwatch.com/2016/03/31/what-if-we-tried-to-replicate-papers-before-theyre-published/>

et échoué partout ailleurs et une dernière a été validée mais ses effets ont été minorés entre l'étude originale et les répliques. Un article a été publié à partir des résultats, rassemblant tous les auteurs impliqués et présentant les avantages et inconvénients de cette méthode de science collaborative (Schweinsberg *et al.*, *forthcoming* voir Figure 1).

Figure 1. Première page de l'article publié à partir d'un projet collaboratif de test de répliquabilité avant publication

The pipeline project: Pre-publication independent replications of a single laboratory's research pipeline

Martin Schweinsberg^{a,*}, Nikhil Madan^a, Michelangelo Vianello^b, S. Amy Sommer^c, Jennifer Jordan^d, Warren Tierney^a, Eli Awtrey^e, Luke Lei Zhu^f, Daniel Diermeier^g, Justin E. Heinze^h, Malavika Srinivasanⁱ, David Tannenbaum^e, Eliza Bivolaru^a, Jason Dana^j, Clinton P. Davis-Stober^k, Christilene du Plessis^l, Quentin F. Gronau^m, Andrew C. Hafenbrackⁿ, Eko Yi Liao^o, Alexander Ly^m, Maarten Marsman^m, Toshio Murase^p, Israr Qureshi^q, Michael Schaefer^a, Nico Thornley^a, Christina M. Tworek^r, Eric-Jan Wagenmakers^m, Lynn Wong^a, Tabitha Anderson^s, Christopher W. Bauman^t, Wendy L. Bedwell^u, Victoria Brescoll^v, Andrew Canavan^s, Jesse J. Chandler^v, Erik Cheries^w, Sapna Cheryan^e, Felix Cheung^{x,y}, Andrei Cimpian^z, Mark A. Clark^z, Diana Cordon^s, Fiery Cushman^l, Peter H. Ditto^l, Thomas Donahue^s, Sarah E. Frick^u, Monica Gamez-Djokic^{a,b}, Rebecca Hofstein Grady^t, Jesse Graham^{ab}, Jun Gu^{ac}, Adam Hahn^{ad}, Brittany E. Hanson^{ae}, Nicole J. Hartwich^{ad}, Kristie Hein^s, Yoel Inbar^{af}, Lily Jiang^e, Tehlyr Kellogg^s, Deanna M. Kennedy^{ag}, Nicole Legate^a, Timo P. Luoma^{ad}, Heidi Maibuecher^s, Peter Meindl^{ah}, Jennifer Miles^l, Alexandra Mislin^z, Daniel C. Molden^{ah}, Matt Motyl^{ae}, George Newman^j, Hoai Huong Ngo^{ah}, Harvey Packham^{ai}, Philip S. Ramsay^u, Jennifer L. Ray^{aj}, Aaron M. Sackett^{ak}, Anne-Laure Sellier^c, Tatiana Sokolova^{ca}, Walter Sowden^h, Daniel Storage^r, Xiaomin Sun^{am}, Jay J. Van Bavel^{aj}, Anthony N. Washburn^{ae}, Cong Wei^{am}, Erik Wetter^{am}, Carlos T. Wilson^s, Sophie-Charlotte Darroux^a, Eric Luis Uhlmann^{a,*}

^a INSEAD, France and Singapore

^b University of Padua, Italy

^c HEC Paris, France

^d University of Groningen, The Netherlands

^e University of Washington, United States

^f University of Manitoba, Canada

^g University of Chicago, United States

^h University of Michigan, United States

ⁱ Harvard University, United States

^j Yale University, United States

^k University of Missouri, United States

^l Rotterdam School of Management, Erasmus University, the Netherlands

^m University of Amsterdam, the Netherlands

ⁿ Carlos Labon School of Business and Economics, Portugal

^o Macau University of Science and Technology, Macau

^p Roosevelt University, United States

^q IE Business School, IE University, Spain

^r University of Illinois at Urbana-Champaign, United States

^s Illinois Institute of Technology, United States

^t University of California, Irvine, United States

^u University of South Florida, United States

^v Mathematica Policy Research, Institute for Social Research, University of Michigan, United States

^w University of Massachusetts Amherst, United States

^x Michigan State University, United States

^y University of Hong Kong, Hong Kong

^z American University, United States

^{aa} Northwestern University, United States

^{ab} University of Southern California, United States

^{ac} Monash University, Australia

^{ad} Social Cognition Center Cologne, University of Cologne, Germany

* Corresponding authors.

E-mail addresses: martin.schweinsberg@insead.edu (M. Schweinsberg), eric.luis.uhlmann@gmail.com (E.L. Uhlmann).

Les avantages théoriques de ce système semblent assez évidents mais il pose des problèmes tout aussi clairs pour les chercheurs qualitatifs. Que signifie la répliquabilité en recherche qualitative ? Et quel en est le coût ? Il semble difficile de concevoir la réplique d'un projet de recherche intervention. Faut-il imaginer d'autres dispositifs de test et d'évaluation de la recherche intervention ? La question demeure ouverte.

Après publication, on trouve le PPPR (*Post publication peer reviewing*) avec des acteurs comme PubPeer. Il s'agit de plateformes de discussions anonymes post-publication des articles par les pairs. Un article suspect, par sa méthodologie par exemple, peut donc être soumis sur la plateforme et discuté par les membres de la communauté. Prenons comme exemple : "Bringing the Institutional Context Back In: A Cross-National Comparison of Alliance Partner Selection and Knowledge Acquisition", de Gurneeta Vasudeva, Jennifer W. Spencer & Hildy J. Teegen publié dans *Organization Science* en 2013. Un commentaire d'un participant non enregistré critique la validité des résultats statistiques (Figure 2), ce qui a provoqué une correction de l'article (voir Figure 3). C'est souvent le cas des articles sur cette plateforme.

0 Unregistered Submission: (May 18th, 2015 7:45pm UTC)

The regression results in this article perhaps require clarification. Table 3 on p. 332 shows regression models with beta coefficients, robust standard errors, and significance tests using the 5%, 1%, and 0.1% levels of significance (one-tailed). This means that t-statistics must be at least 1.64 before a coefficient can be flagged as significant at the lowest threshold of $p < 0.05$. Several of the coefficients do not seem to meet this criterion, even though they are flagged as significant beyond $p = 0.05$, while others do meet this criterion yet are not flagged as significant beyond $p = 0.05$:

Control variables:

- Partners' technological experience (model 1a): $t = 0.04/0.006 = 6.67$ --> not flagged as significant
- Partners' proportion of foreign investors (model 1a): $t = -0.03/0.29 = -0.1$ --> flagged as significant $p < 0.01$
- Partners' portfolio technological impact (model 1a): $t = 28.08192/3.33 = 8.43$ --> not flagged as significant
- No. of prior citations to others in country (model 1a): $t = 0.15/0.15 = 1$ --> flagged as significant $p < 0.01$
- No. of prior alliances between partners (model 1a): $t = 9.82/7.451.77 = 0.001$ --> flagged as significant $p < 0.05$
- Country degree centrality (model 1a): $t = 0.28/0.03 = 9.33$ --> not flagged as significant
- Country degree centrality (model 1d): $t = -0.07/0.05 = -1.4$ --> flagged as significant $p < 0.05$
- Country GDP (model 1c): $t = -0.27/0.19 = -1.42$ --> flagged as significant $p < 0.05$
- Country individualism (model 1a): $t = 0.01/0.02 = 0.5$ --> flagged as significant $p < 0.05$

Explanatory variables:

- Country corporatism * Social value (model 1b): $t = 0.15/0.43 = 0.36$ --> flagged as significant $p < 0.05$
- Country corporatism * Technological value (model 1b): $t = -0.24/0.17 = -1.41$ --> flagged as significant $p < 0.05$
- Country corporatism * Technological value (model 1d): $t = 0.200/0.22 = 0.91$ --> flagged as significant $p < 0.05$

One likely possibility is that the coefficients actually represent standardized regression coefficients (even though in that case, some effect sizes in the table are extremely large, i.e., > 5) and we hope the authors might be able to clarify this.

[Reply](#) [Report](#)

[Permalink](#)

0 Unregistered Submission: (May 27th, 2015 4:17pm UTC)

The table is riddled with inconsistencies. Can the authors perhaps explain the above issues?

Erratum

Permalink: <http://dx.doi.org/10.1287/orsc.2015.0997>
 Published Online: July 31, 2015
 Page Range: 1261 - 1261

[Citation](#) [Full Text](#) [PDF](#) [Related](#)

In the article "Bringing the Institutional Context Back In: A Cross-National Comparison of Alliance Partner Selection and Knowledge Acquisition" by Gurneeta Vasudeva, Jennifer W. Spencer, and Hildy J. Teegen (*Organization Science*, March–April 2013, pp. 319–338), Table 3 on page 332 was incorrect. The table has been corrected in the online version, and the associated explanation of results is provided on page 331 of the online version. The main findings and conclusions of the study remain unchanged.

Figure 2. Extrait de la discussion PubPeer de l'article Vasudeva, et al. 2013

Figure 3. Corrections de l'article sur la page d'Organization Science

Un article du *Journal of Operations Management* a été retiré suite à sa discussion sur PubPeer : "RETRACTED: The relationship between brokers' influence, strength of ties and NPD project outcomes in innovation-driven horizontal networks" Adegoke

Oke, Moronke Idiagbon-Oke, Fred Walumbwa, *Journal of Operations Management* (2008).

PubPeer pose cependant certains problèmes, tels que les fraudes ou le « *sockpuppet* » (faux-nez), le *spam* ou la perturbation des débats, problèmes qui sont ceux des débats en ligne en général. Certes, seuls les premiers et les derniers auteurs d'un article avec un identifiant de type DOI, PubMed ID, ou arXiv ID peuvent créer un compte, et débattre (anonymement ou non), mais les commentaires sont possibles même pour des utilisateurs non enregistrés. Malgré cela, la mise en évidence des failles de certains articles peut conduire à des corrections, des rétractations ou à des accusations de fraude sur des sites comme RetractionWatch, qui opère le nécessaire travail de suivi des rétractations et des dénonciations de fraude. Cependant, pour les revues, le coût d'une rétractation étant beaucoup plus élevé que celui d'une simple correction, la seconde paraît logiquement la plus courante. C'est aussi une des raisons pour lesquelles l'utilisation de la PPIR, présentée précédemment, sa généralisation et sa démocratisation auraient plus d'avantages que la discussion post-publication, même si celle-ci joue un rôle essentiel.

Parmi les devoirs et responsabilités du chercheur vis-à-vis de la communauté académique, on trouve celui de relecture, pour des revues à comité de lecture ou pour des conférences. S'engager dans de tels processus, PPIR, avant publication, et PPPR, post publication, ne font-ils pas autant partie de ces devoirs ?

Repenser la citation : le *superscripting*

Une autre des critiques du système porte sur les indicateurs de publication. Les analyses bibliométriques font le postulat que le nombre de citations d'un papier (et par là, d'une revue) est représentatif de la qualité de la recherche en question. Même ne prendre que les citations de ce papier dans les meilleures revues est intrinsèquement défectueux puisqu'il peut s'agir de « *name dropping* » (le fait de citer sans l'avoir lu une recherche juste parce qu'il est judicieux de la citer). Mais il y a pire : un article peut se retrouver très cité alors qu'il est critiqué et démonté. Les indices de citations reviennent donc à dire que toute publicité est bonne à prendre. C'est le biais de la mesure quantitative indifférenciée.

Anicich (2014) propose donc un système de *superscripting* des bibliographies, *i.e.* un codage des citations en fonction de leur utilisation. Dans l'exemple suivant, ce sont les petits exposants finaux :

Halevy, N., Chou, E. Y., Galinsky, A. D., & Murnighan, J. K. (2012). When hierarchy wins evidence from the national basketball association. *Social Psychological and Personality Science*, 3(4), 398–406.^c

Josephs, R. A., Sellers, J. G., Newman, M. L., & Mehta, P. H. (2006). The mismatch effect: When testosterone and status are at odds. *Journal of Personality and Social Psychology*, 90, 999–1013.^t

Leavitt, H. J. (2005). *Top down: Why hierarchies are here to stay and how to manage them more effectively*. Boston, MA: Harvard Business School Press.^c

Markman, K. D., Lindberg, M. J., Kray, L. J., & Galinsky, A. D. (2007). Implications of counterfactual structure for creative generation and analytical problem-solving. *Personality and Social Psychology Bulletin*, 33, 312–324.^m

Mazur, A., & Booth, A. (1998). Testosterone and dominance in men. *Brain & Behavioral Sciences*, 21, 353–397.^t

Rucker, D. D., & Galinsky, A. D. (2008). Desire to acquire: Powerlessness and compensatory consumption. *Journal of Consumer Research*, 35, 257–267.^M (Anicich, 2014, p. 683)

Anicich propose six catégories de citations : les références dont les résultats sont Consistent (c) avec ceux du papier, sont *Replicated* (r) par le papier, sont *Inconsistent* (i), ou *Failed* (f) lorsque les résultats n'ont pas réussi à être répliqués dans la recherche, (t) lorsque les références ont été utilisées pour construire la *Théorie*, ou (m) lorsqu'elles ont été utilisées d'un point de vue *Méthodologique*. Des indices peuvent être ensuite calculés à partir de ces exposants. Il serait peut-être judicieux de rajouter une catégorie d'auto-citation afin de minorer le calcul des indices.

Ce système a pour avantage de faire peser la majorité de la tâche sur les auteurs eux-mêmes, supposés avoir lu et connaître les références qu'ils utilisent et donc censés pouvoir catégoriser facilement les citations, même si Anicich recommande l'utilisation de tierces parties (les maisons d'édition ?) pour effectuer ce travail afin d'éviter des codages par imitation. L'idée serait de construire des indices de citations plus riches et plus pertinents, mais aussi de suivre la vie des articles, les modes de publication, les tendances dans les journaux et de mettre en évidence les controverses et les écoles de pensée. Si les écueils semblent nombreux et le coût d'entrée trop élevé sans le soutien fort d'institutions et de chercheurs réputés, les réflexions que suscite ce système demeurent intéressantes.

Conclusion

Dans un secteur de plus en plus schizophrénique, qui prétend encourager le collectif tout en idolâtrant les chercheurs *stars* et en évaluant individuellement la recherche ; qui pousse à la publication dans des revues classées dès la thèse tout en réduisant la durée du travail doctoral et en critiquant parfois vertement le résultat final qu'est la thèse par articles ; qui produit une logorrhée académique, superficielle sur les sujets à la mode tout en conservant les classements et autres technologies invisibles qui en sont la cause ; qui se réclame d'un idéal de science ouverte tout en mettant en place des systèmes payants extrêmement lucratifs, les pratiques et dispositifs alternatifs de collaboration ouverte paraissent de plus en plus attractifs de par l'esprit et les valeurs qu'ils véhiculent.

D'aucuns diront peut-être « oui, mais... ». Les satanées technologies invisibles qui structurent les recrutements et ensuite les carrières subsisteront malgré tout. Mais peut-on sincèrement penser que le système actuel est durable, étant donné l'explosion non seulement du volume de publications mais aussi du temps de publication (on peut compter jusqu'à cinq voire huit ans pour une revue de rang A, soulevant ainsi des questions de pertinence des résultats) ? Il ne tient qu'à nous de publier autrement et de favoriser des modes de diffusion du savoir plus diversifiés (blogs, sites d'information comme *TheConversation* ou *Huffington Têtes Chercheuses*, MOOC, et bien sûr livres – encouragés notamment par la FNEGE) mais aussi de militer activement pour un monde de la recherche meilleur. La parcimonie, la joie et le bon goût pourraient être dès lors des valeurs clés du réenchantement ■

Références

Anicich Eric M. (2014) “What Lies Within Superscripting References to Reveal Research Trends”, *Perspectives on Psychological Science*, vol. 9, n° 6, pp. 682-691.

- Berry Michel (1983) *Une technologie invisible ? L'impact des instruments de gestion sur l'évolution des systèmes humains*, Paris, Centre de recherche en Gestion de l'École polytechnique.
- Charroin Jean (2013) "Le classement de Shanghai comme technologie invisible", *Le Libellio d'Aegis*, vol. 9, n° 4, pp. 35-42.
- Kogut Bruce (2008) "Rankings, schools, and final reflections on ideas and taste", *European Management Review*, vol 5, n° 4, pp. 191-194.
- Laroche Hervé (2015) "Sur le professionnalisme dans la recherche", *Le Libellio d'Aegis*, vol. 11, n° 3, pp. 89-93.
- Moisdon Jean-Claude (1997) "Du mode d'existence des outils de gestion", Actes du Séminaire Contradictions et Dynamique des Organisations, Paris, 23 janvier, pp. 5-38.
- Open Science Collaboration (2015) "Estimating the reproducibility of psychological science", *Science*, vol. 349, n° 6251, p. aac4716.
- Osterloh Margit & Frey Bruno S. (2015) "Ranking Games", *Evaluation review*, vol. 39, n° 1, pp. 102-129.
- Osterloh Margit, Frey Bruno S. & Homberg Fabian (2008) "Le chercheur et l'obligation de rendre des comptes", *Gérer et Comprendre*, n° 91 (mars), pp. 48-54.
- Riveline Claude (1991) "Un point de vue d'ingénieur sur la gestion des organisations", *Gérer et Comprendre*, n° 25 (décembre), pp. 50-62.
- Schooler Jonathan W. (2014) "Turning the Lens of Science on Itself Verbal Overshadowing, Replication, and Metascience", *Perspectives on Psychological Science*, vol. 9, n° 5, pp. 579-584.
- Schweinsberg Martin, Madan Nikhil, Vianello Michelangelo, Sommer S. Amy, *et al.* (Forthcoming) "The pipeline project: Pre-publication independent replications of a single laboratory's research pipeline", *Journal of Experimental Social Psychology*.

... pour aller plus loin

- Sci-hub : <http://sci-hub.cc>
- Pubpeer : <https://pubpeer.com>
- Retractionwatch : <http://retractionwatch.com>

