

Références et honnêteté scientifique

Hervé THIS

Les erreurs de citations abondent dans les articles scientifiques, et l'utilisation de l'intelligence artificielle n'est pas une panacée contre cette plaie. Promouvons la connaissance et l'honnêteté.

Mots-clés Publication scientifique, références, erreurs, éthique.

La question des références erronées ou fautives dans les articles scientifiques est si grave qu'il vaut mieux la prendre avec un peu d'humour, sans cesser bien sûr d'être sérieux. D'autre part, produire un article de plus à ce sujet risque d'être inutile, tant la question est claire, avec des règles largement diffusées... qui ne sont pas assez suivies. On tente donc ici un exercice un peu différent, idiosyncratique, moins institutionnel, plus vivant, peut-être un peu « bariolé » apparemment, mais qu'importe, s'il est efficace. Et il doit l'être rapidement, car l'utilisation de l'IA se généralise, ajoutant à la cacophonie éditoriale : littérature grise, systèmes éditoriaux submergés, développement des usines à articles, etc.

Pour examiner la question particulière des citations erronées ou fautives dans les articles scientifiques, nous le faisons de manière moins structurée qu'illustrée, en commençant par des observations récentes (factuelles), avant de rappeler quelques résultats quantitatifs à propos de la plaie des mauvaises références, puis de terminer par une réflexion en vue d'une action de tous les instants.

Des faits, toutefois sans pessimisme

C'est un fait, tout d'abord, que les institutions d'enseignement et les institutions scientifiques, observant la déferlante de l'IA, sont préoccupées par les pratiques déviantes à propos des références dans les articles scientifiques. Mais la crainte n'est-elle pas analogue à celle qui était née, naguère, du développement de *Wikipedia* : pour ces deux cas, une facilité d'usage aurait nuit à la qualité des productions intellectuelles. Observant ainsi que le monde n'a guère changé, il reste toutefois l'espoir que nous pouvons améliorer lentement nos communautés par la diffusion des connaissances.

Mais commençons par des faits, et notamment quelques exemples. Au cours de mes propres études bibliographiques, je vois régulièrement, dans des articles scientifiques qui m'intéressent *a priori*, des références à des articles qui n'ont manifestement pas été lus par ceux qui les citent... puisque je n'y trouve pas ce qui est annoncé. Les auteurs de ces mauvaises citations, qu'ils soient hâtifs, médiocres ou malhonnêtes, sont pris en faute quand on s'aperçoit ultérieurement que les références qu'ils citent ne sont pas justes, qu'on leur écrit pour le leur dire... ou quand on découvre que les articles cités n'existent pas.

J'ai eu le cas tout récemment : intéressé par une citation dans un article, et ne trouvant pas le texte cité, j'ai écrit au dernier auteur (le *corresponding author*, sans doute le patron du laboratoire), qui m'a d'abord répondu qu'il n'avait pas le texte en question, mais qu'il allait le demander au premier auteur (un chercheur postdoctoral). Il y avait là une certaine honnêteté, et, de fait, deux heures plus tard, ce collègue – que j'avais

aussitôt remercié – m'indiquait que son jeune collègue n'avait pas non plus l'article. Quelle honte pour tous les deux !

La fin de l'histoire est édifiante : le premier auteur avait pris une citation dans un article qui était accessible, mais cette citation renvoyait vers une revue scientifique... qui n'a jamais existé. Et quand j'ai fait des recherches à propos de l'auteur de la citation, j'ai fini par découvrir qu'il avait été l'éditeur d'une revue qui avait changé trois fois de nom... et que son article que je cherchais était dans une des trois revues, mais avec un nom de revue qui n'était pas celui qui avait été cité. J'ai alors envoyé l'article que j'avais finalement trouvé à nos deux collègues, à la place de qui je n'aimerais pas être.

Pire est le cas des articles qui sont cités pendant des décennies, en chaîne, sans vérification. Dans les pages de cette revue, j'ai récemment mentionné ce cas à propos des réactions fautivement dites « de Maillard » : depuis les années 1950, des historiens de la chimie ont répété à l'envi la référence à deux auteurs, Ling et Malting... mais la preuve qu'ils n'avaient pas lu le texte correspondant se trouve dans le fait que le brasseur anglais A.R. Ling publia en réalité un article à propos de la confection du malt, *malting* en anglais ; il n'y a eu aucun co-auteur du nom de Malting [1]. La recherche bibliographique sérieuse, facilitée aujourd'hui par l'accès numérique aux sources primaires, doit nous aider à corriger nos erreurs ; autrement dit, nous avons moins d'excuses que par le passé, quand la revue que l'on cherchait se trouvait dans une bibliothèque à l'autre bout du monde. En l'occurrence, à propos des réactions fautivement dites « de Maillard », cette recherche bibliographique sérieuse a établi que ces réactions furent découvertes par le pharmacien lillois Lucien Dusart un demi-siècle avant Maillard ; elles furent ensuite explorées par Hugo Schiff, notamment, puis par toute l'école allemande de chimie, avec notamment Emil Fischer, auquel Louis Camille Maillard se réfère 91 fois dans son mémoire !

J'observe en passant que ces recherches historiques doivent non seulement corriger des références erronées ou fautives, mais, de surcroît, elles doivent nous conduire à changer les dénominations, puisque la règle veut que le nom donné à une découverte revienne au découvreur : Dusart, en l'occurrence. Dans le même ordre d'idée, on évitera de parler d'« émulsion de Pickering » pour désigner des émulsions stabilisées par des particules, car ce type de système a été découvert en 1903, par Ramsden [2]. Pickering n'a étudié ces systèmes qu'en 1907, ayant reconnu l'antériorité de son collègue britannique [3]. Mais, pour s'en rendre compte, il faut faire sa recherche bibliographique sérieusement, afin de bien savoir ce dont on parle : on voit que l'on reste dans le sujet de cet article.

Des faits toujours : il n'est pas éthique de citer des textes secondaires, parce que l'on prive alors les vrais auteurs de la paternité de leur découverte, de leur propriété intellectuelle

Pour la science et la technologie des aliments (avec souvent plus de technologie que de science), le cas est plus que quotidien, surtout depuis que se multiplient des synthèses bibliographiques. D'ailleurs, il y a pire : les revues indignes qui publient les articles dont les références sont inappropriées vont parfois jusqu'à réclamer que les références aient moins de cinq ans : là, c'est une incitation à la malhonnêteté... et, au fond, pourquoi ces revues ont-elles cette pratique ? Du temps où le papier coûtait aux éditeurs, on comprenait qu'il fallait limiter le volume des articles (et en refuser beaucoup, pour les mêmes raisons), mais aujourd'hui, alors que les publications sont électroniques, pourquoi cette limitation ?

Des études expérimentales pour établir les faits dans leur généralité

On pourrait penser que les exemples du paragraphe précédent sont anecdotiques... mais ce serait ignorer que, au fil des années, des études expérimentales ont exploré quantitativement la question des références dans les articles scientifiques, et qu'elles ont parfaitement établi la grande généralité de tout cela.

Par exemple, en 2002, Simkin et Roychowdhury ont mis en œuvre une méthode statistique pour estimer la proportion de scientifiques qui ont effectivement lu les articles qu'ils citent [4], en étudiant une loi empirique (la loi de Zipf) qui décrit la répartition des erreurs dans les citations [5]. En étant charitables, ils concluent que seulement 20 % des articles cités ont été vraiment lus entièrement.

Puis en 2020, Smith et Cumberledge ont repris la question, examinant 250 citations prises au hasard, et les comparant soigneusement aux textes cités : 25 % de ce qui était prétendu cité ne l'était pas [6] !

En 2021, Pavlovic *et coll.* se sont d'abord focalisés sur la précision des citations dans 1540 articles de revues biomédicales (2526 citations), avec le concours des premiers auteurs [7]. Puis les vérifications ont porté sur 4912 citations dans 2995 articles : il est apparu qu'il y avait au moins une citation imprécise dans respectivement 11 et 15 % des articles du premier et deuxième groupe. Dans 38,4 % des cas, les résultats cités n'existaient pas ; dans 15,4 % des cas, les résultats cités n'étaient pas correctement interprétés, et un cinquième des citations erronées résultait de citations imprécises en chaîne. Plus anecdotique, une étude de 2023 a porté sur un manuscrit soumis à la revue *World Journal of Mens Health*, avec 172 références à 145 articles : les rapporteurs ont dépisté des erreurs dans 20 % des citations ; neuf citations étaient improprement données, six extrapolaient abusivement des résultats cités, quatre références étaient secondaires, avec deux vers des références qui n'existaient pas, par exemple [8].

Mais évidemment, l'emploi de méthodes numériques permet de bien mieux cerner ce grave problème des références : en 2025, Baethge et Jergas ont fait une étude systématique des références (32 000) dans *PubMed* et *Web of Science*, ainsi qu'une méta-analyse de telles études (46), pour des textes médicaux [9]. Il est apparu que 16,9 % des citations étaient incorrectes, avec environ la moitié d'erreurs « majeures », et sans améliorations dans les dernières années, malgré la facilité accrue de citer correctement. Les auteurs faisaient la différence entre des imprécisions de citation (des noms d'auteurs mal orthographiés, des erreurs dans les numéros de pages) et des erreurs « majeures », de contenu : les résultats qu'ils ont publiés ne concernaient que ces dernières.

Pourquoi ?

J'écris tout cela sans prétendre être blanc : je dois avouer que, il y a bien longtemps, cherchant pourtant à bien faire, je me faisais une règle de citer des articles récents, montrant que j'étais au fait des publications scientifiques. C'était un péché par ignorance : je n'avais pas compris que je privais ainsi de leur paternité, de leur « propriété intellectuelle », les auteurs des découvertes ou des idées auxquelles je me référais. J'ignorais ainsi la règle éthique notamment – et justement – édictée par le *Committee of Publication Ethics* (et par bien d'autres institutions), ce qui, avec les années, est devenu d'autant plus grave que l'évaluation des chercheurs s'est faite sur les citations [10,11].

Comment bien citer, alors ? C'est très simple : s'il y a une idée ou un résultat expérimental, on cite l'auteur de l'idée ou de la découverte. Et s'il y a eu des amendements ultérieurs de cette idée ou de cette découverte, on cite les auteurs initiaux et les auteurs des amendements. Et c'est ainsi que l'on est conduit à citer une chaîne continue de progrès.

Tout cela survient alors que l'intelligence artificielle est de plus en plus utilisée, et nos amis feraient bien de se méfier : plusieurs fois de suite (et avec une proportion très élevée), récemment, l'IA m'a donné des réponses fausses, et, quand je le lui signalais après vérification, elle m'a répondu en substance : vous avez raison [elle cite alors ce que je viens de lui dire], et elle a embelli... avec d'autres informations dont je doute encore.

Pour sourire un peu, alors que les choses sont pourtant graves, ce cas m'est encore venu il y a quelques minutes, quand, demandant à ChatGpt qui était le premier auteur à avoir discuté explicitement l'éthique des citations dans les articles scientifiques, l'IA m'a répondu en me donnant le nom d'un chirurgien anglais ; quand j'ai demandé la référence du texte où la question aurait été traitée, j'ai reçu la réponse : « *Je dois corriger ce que j'ai dit précédemment : je ne retrouve pas de texte précis de William Alanson White contenant la critique explicite des auteurs citant des travaux qu'ils n'ont pas lus. Les recherches disponibles ne permettent pas d'attribuer solidement à White cette remarque* » [sic].

On sait bien que des scientifiques malhonnêtes (méritent-ils d'être dits scientifiques ?) auront recours à des IA pour produire des textes de mauvaise qualité (on voit que je ne dis pas ici que seraient malhonnêtes tous ceux qui utilisent l'IA, mais seulement ceux qui l'utilisent à mauvais escient), mais la question est moins de dépister ces articles mauvais que de bien expliquer les règles de l'honnêteté, et, surtout, de supprimer les mécanismes qui incitent des malhonnêtes à publier de mauvais articles, avec notamment de mauvaises citations. Notamment, dans certaines institutions, on a jugé, et l'on juge encore, les chercheurs au nombre d'articles publiés : c'est d'autant plus mauvais que prolifèrent des journaux « gris » ou « noirs », où les auteurs payent pour être publiés, ce qui est presque une garantie de mauvaise qualité. Il faut certainement dénoncer de telles revues, en publier la liste, pour que soit patent que ceux qui y publient ne sont guère honnêtes... mais cela ne règle pas la question des références erronées ou fautives.

Comment éviter les mauvais comportements ?

Surtout, les constats doivent être suivis d'actions. Et si les lois n'ont jamais arrêté les malhonnêtes, il y a lieu de combattre

d'abord les ignorances, par l'éducation, par l'instruction, par l'enseignement. On citera donc, parmi d'autres, Penders, qui donne dix règles pour citer correctement, après avoir expliqué comment les citations non seulement indiquent comment nos propres travaux se situent dans la chaîne du savoir scientifique, mais aussi comment elles nous confient une responsabilité politique et morale, rappelant que les citations servent à évaluer les travaux et les chercheurs, à distribuer les crédits de recherche, par exemple [12].

Parmi les dix règles, il y a d'abord celle qui consiste à bien inclure toutes les références à des articles importants pour situer son propre travail. Puis il faut citer les publications que l'on mentionne, et les citer d'ailleurs en relation avec le contenu, sans focaliser l'attention des lecteurs de nos manuscrits sur des détails. Il faut citer avec transparence, mais pas avec neutralité : les citations doivent donner une évaluation, signaler un intérêt particulier, une importance particulière dans le contexte de nos propres travaux. Il faut aussi se citer... quand c'est nécessaire, et seulement dans ce cas. Et il faut prioriser les citations que l'on conserve finalement, surtout quand le nombre a été limité par l'éditeur.

Et ainsi de suite : je renvoie vers l'article de Penders... qui mérite non seulement d'être cité, mais d'être lu. À faire étudier dès le premier cycle des universités ? Car c'est de l'« école » à l'« université », et au-delà, qu'il faut diffuser des idées justes, des explications, afin que « la bonne monnaie chasse la mauvaise ».

Surtout, je propose de moins manier le bâton que la carotte. Par exemple en levant les yeux, de la fange vers le ciel bleu : contemplons les personnalités exemplaires, qui nous ont précédés. N'oublions pas de faire rêver : si l'hagiographie a été souvent – et justement – récusée par les historiens des sciences, il reste quelques beaux exemples que nous pouvons tendre sans fard à nos jeunes collègues et à nous-mêmes.

Ainsi, en 1918, dans son discours prononcé pour le 60^e anniversaire de Max Planck, Albert Einstein avait évoqué un « temple de la science », d'où un ange chasserait tous ceux qui ne font pas de recherche scientifique assez « purement » : pour de l'argent, pour du pouvoir, pour du sport... Qui resterait alors ? Ceux qui cherchent naïvement et honnêtement à lever un coin du grand voile, émerveillés par le mystère du monde [13].

Einstein venait d'ailleurs, de ce point de vue, après Michael Faraday, sans oublier tous les autres, moins notoires, tous ceux qui cherchent à faire bien, à faire mieux, et, évidemment, qui s'étonnent des comportements médiocres... comme le sont aujourd'hui bien des pratiques de citation des références bibliographiques. Pour Faraday, que les hasards de la naissance semblaient destiner au métier de relieur, il écrivit : « *Mon désir d'échapper au monde du commerce, qui me paraissait vicieux et égoïste, et d'entrer au service de la Science, dont je m'imaginais qu'elle rendait ceux qui la cultivent aimables et généreux, me conduisit finalement à entreprendre la démarche simple et audacieuse d'écrire à Sir H. Davy [...]; je lui adressais en même temps les notes que j'avais prises de ses conférences* » [14]. Et c'est ainsi que Faraday finit par devenir l'un des plus grands physico-chimistes de son siècle, loué par tous, mais, en particulier, par le chimiste Jean-Baptiste Dumas : « *Un professeur plein de feu pour la diffusion de la vérité, un artiste infatigable,*

plein d'entrain et de gaieté dans son laboratoire, le meilleur et le plus doux des hommes au sein de sa famille, et le prédicateur le mieux inspiré au milieu de l'humble troupeau religieux dont il suivait la foi. La simplicité de son cœur, sa candeur, son amour ardent de la vérité, sa franche sympathie pour tous les succès, son admiration naïve pour les découvertes d'autrui, sa modestie naturelle, dès qu'il s'agissait des siennes, son âme noble, indépendante et fière, tout cet ensemble donnait un charme incomparable à la physionomie de l'illustre physicien. Je n'ai pas connu d'homme plus digne d'être aimé ; d'être admiré ; d'être regretté. » Avec Faraday, les qualités morales étaient garantes des qualités scientifiques : exactitude, honnêteté, persévérance, etc. La modestie de caractère était associée à la modestie des moyens nécessaires à la recherche scientifique. Montrons son exemple à tous... pour que s'améliorent les statistiques en matière de références dans les articles scientifiques.

[1] H. This vo Kientza, Why we should speak of glycation reactions, or of amino-carbonyl reactions, instead of Maillard reactions, *MARS Highlights*, **2023**, 18(3), p. 5-19, www.imarsonline.com

[2] W. Ramsden, Separation of solids in the surface-layers of solutions and "suspensions" (observations on surface-membranes, bubbles, emulsions, and mechanical coagulation), Preliminary account, *Proc. R. Soc. Lond.*, **1904**, 72, p. 56-164.

[3] S. U. Pickering, Emulsions, *J. Chem. Soc.*, **1907**, 91, p. 2001-21, <https://pubs.rsc.org/en/content/articlelanding/1907/ct/ct9079102001>

[4] M. V. Simkin, V. P. Roychowdhury, Read before you cite!, *Complex Systems*, **2003**, 14, p. 269-274, <https://doi.org/10.25088/ComplexSystems.14.3.269>

[5] G.K. Zipf, Selected Studies of the Principle of Relative Frequency in Language, Cambridge (Mass.), Oxford University Press, **1932**, https://pure.mpg.de/rest/items/item_2407800_3/component/file_2459540/content

[6] N. Smith, A. Cumberledge, Quotation errors in general science journals, *Proc. R. Soc. A*, **2020**, 476, 20200538, <http://dx.doi.org/10.1098/rspa.2020.0538>

[7] V. Pavlovic et al., How accurate are citations of frequently cited papers in biomedical literature?, *Clin. Sci. (Lond)*, **2021**, 135(5) p. 671-681, <http://doi.org/10.1042/CS20201573>

[8] A. Agarwal et al., Citation Errors in Scientific Research and Publications: Causes, Consequences, and Remedies, *World J. Mens. Health*, **2023**, 41(3), p. 461-465, <https://doi.org/10.5534/wjmh.230001>

[9] C. Baethge, H. Jergas, Systematic review and meta-analysis of quotation inaccuracy in medicine, *Res. Integr. Peer Rev.*, **2025**, 10, 13, <https://doi.org/10.1186/s41073-025-00173-z>

[10] COPE, **2026**, <https://publicationethics.org/search?query=citation> (consulté le 01/06/2026).

[11] ORI, Ethically Questionable Citation Practices, **2023**, <https://ori.hhs.gov/ethically-questionable-citation-practices> (consulté le 01/06/2026).

[12] B. Penders, Ten simple rules for responsible referencing, *PLoS Comput. Biol.*, **2018**, 14(4), e1006036, <https://doi.org/10.1371/journal.pcbi.1006036>

[13] A. Einstein, Motivation for Exploration (a speech delivered at the celebration of Planck's Sixtieth Birthday in Berlin), **1918**.

[14] J.B. Dumas, Éloge historique de Michel Faraday, Institut impérial de France, Académie des sciences, Firmin Didot, Paris, **1868**, <https://catalogue.bnf.fr/ark:/12148/cb30373835x>

Hervé THIS* est physico-chimiste, directeur du Centre international de gastronomie moléculaire et physique, INRAE-AgroParisTech, Palaiseau.

* herve.this@agroparistech.fr