

Peut-on faire mieux que le *peer review* ?

Une enquête menée auprès des physiciens et physiciennes en juin 2018, par la Commission des publications de la Société Française de Physique

Les publications scientifiques sont la carte de visite des chercheurs, elles traduisent leurs découvertes et avancées, font connaître leurs travaux et sont un moyen important de communiquer les progrès de leur recherche. Par ailleurs, elles sont aujourd'hui la base du processus d'évaluation, celle d'une personne, d'un projet ou d'une unité de recherche.

Contexte de l'enquête

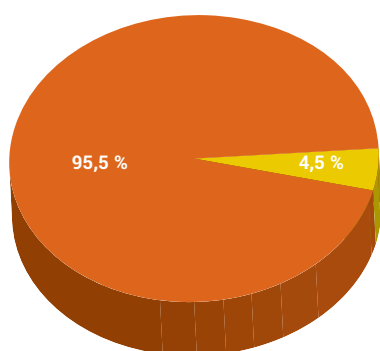
Ce n'est qu'une hypothèse, mais elle est considérée valide et juste par une grande majorité de physiciens : la publication d'un article scientifique (en physique) ne prend sa valeur complète que si elle passe par un processus d'évaluation. Aujourd'hui, dans la plupart des cas, ce processus est une relecture par des pairs (*peer review*), très souvent dans sa forme « simple – aveugle », c'est-à-dire que l'éditeur scientifique et les relecteurs (ou rapporteurs, *reviewers*) connaissent le nom de l'auteur, qui, lui, ne connaît que le nom de l'éditeur scientifique. Ce mécanisme est souvent critiqué : le relecteur peut avoir une lecture biaisée par inadvertance ou par mauvaise volonté, être trop subjectif, pas toujours expert ou trop lent. Il peut favoriser les équipes et chercheurs qui lui sont proches ou très connus ou, pire encore, freiner la publication de vraies découvertes venant de la concurrence. Chacun connaît une anecdote à ce sujet et, même sans ces « faits divers », on sait que parfois la publication originale et pertinente dont on était si fier ne passe que moyennant d'âpres luttes avec certains relecteurs, tandis qu'une pièce cousue un peu rapidement trouve grâce auprès du premier évaluateur. En revanche, devant le nombre croissant de publications, en particulier en provenance des pays du BRIC^(a), on constate aujourd'hui que les sollicitations à expertiser des articles se multiplient pour tous les chercheurs. Ceux d'entre nous qui sont impliqués en tant qu'éditeur scientifique se rendent compte qu'il est de plus en plus difficile de trouver de bons rapporteurs. Alors se pose la question : le *peer review* est-il devenu obsolète ? Est-il à bout de souffle ? Doit-on le réinventer ou l'abandonner ?

Afin de prendre le pouls de la communauté scientifique sur ce sujet, nous avons recueilli les avis des membres de la Société Française de Physique dans une enquête menée en juin 2018. Une vingtaine de questions dans un formulaire en ligne ont recueilli 155 réponses. Pour presque toutes les questions, nous avons laissé de la place pour des commentaires, possibilité largement saisie par les sondés, que l'on remercie vivement pour leur implication dans le débat. L'enquête a été complétée par un test comparant diverses approches du *peer review* sur un article scientifique (voir l'encadré, p. 53). En août 2018, lors des Journées de la Matière Condensée (JMC) à Grenoble, un débat a eu lieu pour conclure l'enquête.

.....
La majorité des sondés s'accorde pour penser que le peer review est une étape essentielle à maintenir dans la publication de nos articles, mais également qu'il nécessite quelques innovations pour s'adapter aux comportements des différentes communautés.

La reconnaissance de l'activité du rapporteur reste un problème majeur si l'on veut maintenir ce gage de qualité.

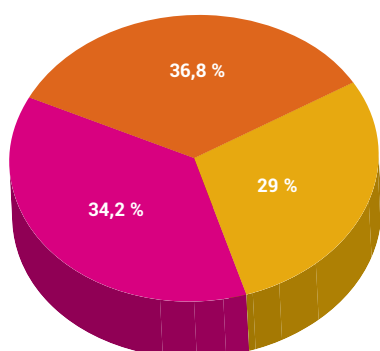
.....



Pensez-vous que l'évaluation de la qualité d'un article scientifique, quelle que soit sa forme, est importante avant sa publication ?

■ Oui : 95,5 % ■ Non : 4,5 %

1. La très grande majorité des sondés pense qu'un processus de *peer review* est important.



Le *peer review* peut être biaisé par la compétition entre équipes, une liste d'auteurs prestigieuse, ou d'autres facteurs. Pour éviter ces biais, il faudrait que le *peer review* :

- Soit fait en double aveugle (le relecteur ne voit pas le nom de l'auteur) : 36,8 %
- Soit fait en triple aveugle (ni le relecteur ni l'éditeur scientifique ne connaissent le nom de l'auteur) : 29 %
- Soit au contraire un processus transparent où les noms des relecteurs sont publiés : 34,2 %

2. Différentes pistes existent pour s'affranchir des biais du *peer review*. Aucun des trois processus ne convainc une majorité de physiciens.

Profil des personnes qui ont répondu

Il s'agissait pour deux-tiers d'enseignants-chercheurs et de chercheurs en activité, le tiers restant étant largement dominé par des collègues retraités ou émérites (15%) et des postdoctorants (10%). Il faut noter que seules 11% des réponses émanaient de femmes. Très peu ont indiqué qu'ils étaient doctorant, ingénieur ou physicien dans l'industrie. Le rapport du nombre de théoriciens au nombre d'expérimentateurs était de deux sur trois. Les questions sollicitaient l'avis de nos collègues auteurs et/ou relecteurs et, dans la plupart des cas, plusieurs réponses étaient possibles.

Valeur et intérêt du *peer review*

À notre première question « Pensez-vous que l'évaluation de la qualité d'un article scientifique (rigueur, méthode employée, accessibilité du texte, intérêt pour la communauté scientifique...), quelle que soit sa forme, est importante avant sa publication ? », un « oui » écrasant à plus de 95% est un indicateur fort (fig. 1). Les commentaires étaient nombreux et portaient majoritairement sur la vérification et la validation du contenu scientifique, sur la limitation nécessaire du nombre de (mauvais) articles, mais aussi sur l'enrichissement du contenu d'un article grâce à un regard extérieur.

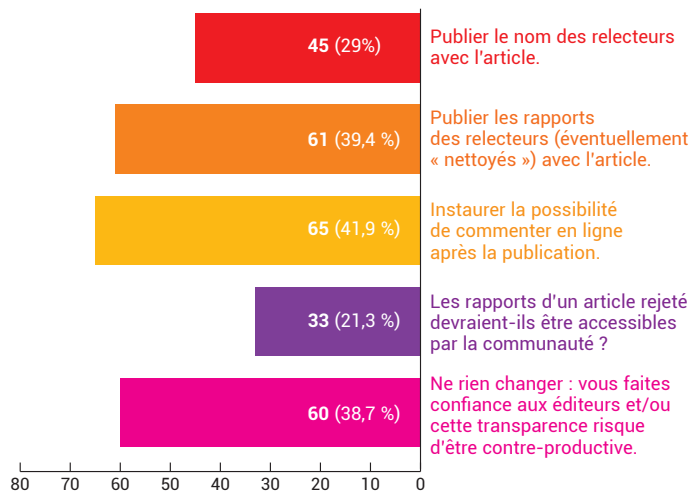
Dans le contexte contemporain de «likes» et de «retweets» pour s'exprimer dans le monde numérique, nous voulions savoir si d'autres formes d'évaluation étaient plébiscitées. À notre grande surprise, très peu de personnes (< 10%) estiment que la certification de qualité peut se faire, soit par cooptation par une communauté soit par statistiques bibliométriques *a posteriori*.

Avis des relecteurs

Le processus actuel de *peer review* est lent : relire et comprendre un article prend beaucoup de temps. On notera que 60% des collègues ont indiqué qu'ils acceptaient de relire les articles dans d'autres thématiques que la leur, parce que cela leur permettait de vraiment se plonger dans un travail qu'ils n'auraient pas forcément lu. Seulement 7% des relecteurs ont eu l'impression d'avoir perdu leur temps devant le refus des auteurs ou de l'éditeur de prendre en compte leurs remarques.

Certaines expérimentations pour faire évoluer le *peer review* sont intéressantes, mais les communautés réagissent différemment. Cela explique sans doute pourquoi, à la question « triple aveugle, double aveugle, ou chacun connaît chacun ? », un équilibre se manifeste entre les trois options (fig. 2). S'il n'est certes pas utile de proposer des relectures en double aveugle dans une petite

communauté expérimentale, le processus de relecture pourrait être amélioré et devenir plus transparent. La figure 3 montre les réponses à la question sur l'amélioration de la transparence du *peer review*. Plusieurs pistes sont envisagées ; leur mise en œuvre dépendra certainement de la taille et des réactions de la communauté, voire de la revue concernée.



3. Quelques solutions envisagées pour rendre le processus de *peer review* plus transparent. Les réponses ne sont pas unanimes ; elles reflètent entre autres l'hétérogénéité des sous-disciplines (plusieurs réponses étaient possibles).

Reconnaitre le travail d'évaluateur

Pour 60,6 % des collègues, la mission du relecteur est de donner un avis sur le contenu scientifique, alors que 36,8 % estiment qu'un avis sur le format technique est aussi souhaitable. L'intérêt de l'article pour la communauté au sens large a visiblement peu d'importance ! Les relecteurs ont une perception très positive de leur travail. Pas moins de 87% d'entre eux trouvent que le manuscrit a été amélioré après leurs commentaires. Cet avis est d'ailleurs partagé par une grande majorité des auteurs.

Se pose alors la question de la reconnaissance et de la valorisation du travail des relecteurs (temps investi, valeur ajoutée...), travail invisible mais qui a un rôle crucial. Cette question est d'autant plus importante qu'avec le nombre croissant des articles soumis, la charge de relecture devient de plus en plus lourde. Presqu'un tiers des collègues sont sollicités deux fois par mois ou plus pour ces expertises. Il y a aujourd'hui des approches variées : certaines maisons d'édition décernent des titres ou des citations honorifiques aux *referees*, d'autres offrent des livres ou des réductions sur les frais de publications (à valoir pour une prochaine publication). Il est évident que le *peer review* fait partie de nos activités de chercheur, et qu'il doit être reconnu comme tel. À la question « Quelle pourrait être la meilleure reconnaissance du *referee* ? », presque deux tiers des réponses ont insisté sur « une inclusion du nombre de vos rapports et le nom des journaux dans votre CV et une reconnaissance de ce travail par les instances d'évaluation des chercheurs ». Vu la confidentialité du processus, la façon dont ceci peut être mis en œuvre n'est pas claire. Plusieurs initiatives ont vu le jour, mais elles sont l'apanage de grandes maisons d'édition, comme Publons [1] ou My Elsevier Reviews Profile [2].

Avis des auteurs

Une grande majorité de relecteurs et d'auteurs sont du même avis. C'est surtout la question de la valeur ajoutée du *peer review* qui est très largement partagée, puisque 89% des auteurs pensent que « la qualité de mon manuscrit s'est souvent trouvée améliorée après les rapports ». Un petit pourcentage d'auteurs constate que leur travail n'a pas été bien compris par les relecteurs, ce qui montre la limite du système, puisque pour des raisons d'éthique le rapporteur consulté par l'éditeur scientifique ne peut pas être trop proche de l'auteur. Cette réaction montre également que la mission du relecteur n'est pas seulement de valider ou non le contenu scientifique, mais aussi d'être le premier « cobaye » externe sur la qualité pédagogique et la clarté d'un article. Par ailleurs, quelques auteurs pointent la lenteur du processus de relecture, qui retarde la publication de leurs résultats. Nous ne pouvons qu'encourager ces collègues à partager leurs *preprints* dès la soumission dans une archive ouverte comme HAL ou arXiv afin de faire connaître leurs travaux au plus tôt, et de rajouter la référence bibliographique une fois l'article validé et publié dans sa version définitive.

Presque tous les auteurs sont d'avis que le *peer review* doit être organisé par la revue, soit par un scientifique engagé comme éditeur (72,9%), ce qui est effectivement le cas pour la plupart des revues des sociétés savantes (*EPL* et *EPJ* pour la SFP), soit par un éditeur scientifique professionnel (21,3%) comme c'est le cas par exemple pour l'American Physical Society ou Nature, mais que ce n'est pas le rôle d'un serveur de *preprint* comme ArXiv. Pour compléter le débat, il y a des réflexions sur la « portabilité » du rapport de relecture permettant à l'auteur, en cas de rejet par un journal, de resoumettre son manuscrit accompagné du premier rapport à une autre revue, ce qui peut accélérer le processus d'évaluation. Aujourd'hui, ceci est pratiqué au sein d'une même maison d'édition ou entre des revues voisines (comme *EPL*, *EPJ* et *J. Phys.*), mais un partage plus large pourrait être une solution à l'avenir.

Le débat avec la communauté des Journées de la Matière Condensée

Les résultats préliminaires de cette enquête ont été présentés et débattus en août 2018 aux JMC de Grenoble, avec une audience d'une centaine de participants. La communauté s'est montrée extrêmement intéressée par ces questions, qui touchent notre travail quotidien de chercheur. Malgré ses faiblesses, le *peer review* est perçu comme étant le meilleur processus de certification de la qualité d'un article avant sa publication. En tant que physicien, on ne fait confiance qu'à nos pairs pour juger la qualité de nos écrits. Mieux, de nombreux collègues ont dit que – sachant qu'il y aurait relecture par un pair – ils faisaient un effort de clarification et d'explication dès la rédaction du premier brouillon d'un article. En effet, le *peer review* est un filtre de qualité qui permet de stabiliser et de limiter la taille d'un texte et à chacun d'accorder une certaine confiance à la qualité des résultats publiés.

Conclusion

Les résultats de notre enquête et le débat qui a suivi, sans être exhaustif, nous démontrent que la communauté de la SFP se sent très impliquée dans le débat sur le *peer review* dans les journaux, et prend le rôle de rapporteur très au sérieux.

La majorité des sondés s'accorde sur le fait que le *peer review* est une étape essentielle à maintenir dans la publication de nos articles, mais également qu'il nécessite quelques innovations pour s'adapter aux comportements des différentes communautés. Il est évident que le *peer review* joue un rôle capital dans l'amélioration des manuscrits, et que le rapporteur doit se pencher en priorité sur la qualité scientifique d'un article, sans pour autant oublier son format technique.

La reconnaissance de l'activité du rapporteur reste un problème majeur si l'on veut maintenir ce gage de qualité ; il doit donc être mieux considéré par les instances d'évaluation des carrières.

Tout en gardant le meilleur de ce système précieux qu'est le *peer review*, mis en place dès la création des premières revues scientifiques au XVII^e siècle, le *peer review* du XXI^e siècle doit être plus transparent dans son processus et, surtout, trouver un moyen de valoriser correctement ce travail fondamental pour la science. ■

Martina Knoop⁽¹⁾ (martina.knoop@univ-amu.fr),
Agnès Henri⁽²⁾, **Jean Daillant**⁽³⁾ et **Bart van Tiggelen**⁽⁴⁾

(1) PIIM, CNRS/Aix-Marseille Université, 13013 Marseille

(2) EDP Sciences, 91944 Les Ulis

(3) Synchrotron SOLEIL, 91190 Saint-Aubin

(4) LPMMC, Université Grenoble Alpes/CNRS, 38042 Grenoble

(a) L'acronyme BRIC se réfère à quatre pays, le Brésil, la Russie, l'Inde et la Chine, qui sont en développement économique rapide depuis le début du 21^e siècle.

1 • <https://publons.com> . Publons permet à un chercheur de se créer un profil sur lequel il peut valoriser les travaux d'évaluation qu'il a fait. Publons, initiative indépendante au départ, a été récemment racheté par Clarivate, le propriétaire de *Web of Science*.

2 • www.reviewerrecognition.elsevier.com/#/ .

Un test en *live* sur le *peer review*

Pour aller plus loin, nous avons souhaité faire un test avec quelques volontaires recrutés au cours de l'enquête. Le panel était petit (7 personnes), bien que plus grand que le nombre de rapporteurs habituels (2), et venait de différentes sous-disciplines de la physique (matière condensée, physique statistique, matière molle, optique). L'objectif du test était de comparer les différentes approches du *peer review* pour voir leur impact sur les résultats et tester la validité de l'évaluation par les pairs.

Le manuscrit testé est la première version d'un article de l'un d'entre nous, portant sur un sujet principalement de matière molle, et qui a été publié, après d'importantes modifications, dans *Phys. Rev. Lett.* **116** (2016) 228101.

Cet article a été soumis aux volontaires en leur demandant d'indiquer s'il était prêt pour publication dans le format proposé, de détailler ses points faibles et forts, et de se prononcer sur les futures citations du papier dans la littérature scientifique. La dernière question était : « Si cet article était soumis à *Nature* ou *Science*, votre jugement serait-il identique ? »

La réaction la plus forte a été la plus révélatrice. Nous avons envoyé le manuscrit seulement quelques jours avant la date limite à un panel initialement plus grand, et une bonne moitié des collègues ont (parfois de façon assez virulente) exprimé leur désaccord sur le délai imposé. Ils étaient unanimes à dire qu'une relecture sérieuse demande du temps. Qu'ils soient rassurés, nous sommes évidemment d'accord !

Il est intéressant de voir par la suite que la différence d'expertise sur le sujet n'avait presque pas d'influence sur les commentaires des relecteurs. La plupart d'entre eux se sont prononcés contre la publication en l'état du texte, en indiquant comme principale faiblesse une présentation trop spécialisée et une argumentation inaccessible à une large audience de physiciens. Sans surprise, les experts plus proches du sujet ont commenté certains détails des travaux, qui n'étaient pas forcément la première cause de la révision demandée. Quant aux forces du manuscrit, tous les relecteurs ont émis un avis positif sur cette présentation de la validation expérimentale d'une prédiction théorique. Ils étaient également unanimes à dire que les citations futures proviendraient surtout des experts du domaine. Enfin, la moitié des rapporteurs auraient rejeté le manuscrit dans le cadre d'une évaluation pour *Nature* ou *Science*, avec l'argument d'un manque d'intérêt pour une communauté plus large.

Il est intéressant de noter que tous les points mentionnés par notre panel ont été soulevés par les rapporteurs initiaux lors de la soumission dans la revue. Aussi, l'avis du panel très hétérogène sur le contenu du manuscrit, ses forces et ses faiblesses, était très similaire aux réactions des rapporteurs de la revue. Nous sommes donc tentés de conclure qu'il existe une certaine cohérence dans l'évaluation, malgré l'hétérogénéité des rapporteurs. Par ailleurs, nous fûmes surpris de constater que le *peer review* pour des journaux de prestige comme *Nature* et *Science* ne suit pas exactement les critères habituels. Ceci est d'autant plus étonnant que les chercheurs reprochent souvent à ces journaux leur ligne éditoriale, avec un processus de *peer review* trop souvent biaisé, voire absent. Est-ce qu'un rapporteur devrait s'exprimer sur le niveau de « prestige » d'un article X pour être compatible avec la « réputation prestigieuse » de la revue Y ? Une belle question pour un futur débat !