
N° 9 | 2020

Les enjeux et les défis de la fonction d'évaluation en sciences de l'information et de la communication

Modes d'évaluation ouverte par les pairs : de la revue à la plateforme

Evelyne BROUDOUX *Maître de conférences HDR*

DICEN

Conservatoire National des Arts et Métiers, Paris

Madjid IHADJADENE

Édition électronique :

URL :

<https://revue-cossi.numerev.com/articles/revue-9/16-modes-d-evaluation-ouverte-par-les-pairs-de-la-revue-a-la-plateforme>

DOI : numerev_1628

Date de publication : 30/11/2020

Cette publication est sous licence **CC BY-NC-ND** (Attribution - No commercial - No derivatives).

Pour **citer cette publication** : BROUDOUX, E., IHADJADENE, M. (2020) Modes d'évaluation ouverte par les pairs : de la revue à la plateforme. *Revue COSSI*, (9). https://doi.org/10.34745/numerev_1628

Cet article a pour but de proposer un état de l'art des différentes formes de l'évaluation d'articles ou de communications par les pairs. De l'évaluation « aveugle » à l'évaluation « ouverte », de multiples possibilités existent et sont expérimentées. C'est dans le champ des sciences que l'on trouve le plus d'innovations sociotechniques s'appuyant sur des plateformes de publication modélisant des workflows éditoriaux originaux. L'ouverture de l'évaluation peut se produire entre pairs, en rendant publiques les identités et/ou les rapports des évaluateurs, à différents stades de l'article scientifique : préprint, en cours de rédaction, ou encore après publication. Cet état de l'art est basé sur un ensemble de publications essentiellement produites par les acteurs de l'évaluation ouverte, issus principalement des disciplines STM.

Mots-clés :

Open Science, Science ouverte, Médiatisation, Évaluation par les pairs, Évaluation ouverte

Abstract : This article wants to provide a state of the art of different forms of peer-reviewed scholarly communication. From masked assessment to open peer-review, multiple possibilities exist and are experienced. It is in the field of science that we find the most sociotechnical innovations based on publishing platforms modeling original editorial workflows. The opening of the evaluation can occur among peers, by making public the identities and / or reports of the evaluators, at different stages of the making of scientific article: preprint, in draft, or after publication. This state of art is based on a set of publications produced by open peer-review actors, mainly from STM disciplines.

Keywords : Peer-review, open peer-review, open science, mediatization

INTRODUCTION

Face à la fonction d'évaluation devenue un mode de gestion des activités en entreprise dans des dispositifs de contrôle et de recherche de performance, il existe un secteur où des procédures remplissent des objectifs de tri, de sélection et d'amélioration. C'est celui de la recherche scientifique où « l'évaluation par les pairs » est au coeur des pratiques de publication. Nous nous intéressons ici aux parties prenantes de ce processus qui tend aujourd'hui à se modifier dans une volonté d'ouverture.

On peut faire remonter le processus d'évaluation de la production scientifique par les

pairs à la fondation des sciences telles que nous les expérimentons encore aujourd'hui. Dans un article sur les acteurs de l'écosystème de la communication scientifique, Jean-Claude Guéron (2001) revenait sur la création en 1665 de la revue « *Phil Trans* », la première à sélectionner pour les diffuser un ensemble de contributions originales porteuses de connaissances[1]. Guéron remarque qu'elle assurait aussi une autre fonction, celle d'être un registre public permettant le repérage de la paternité des œuvres, au moment de la lutte naissante pour la reconnaissance de la propriété intellectuelle au sein du marché émergent de l'imprimerie-librairie. Ceci nous apparaît comme une preuve de la liaison historique entre la fonction d'évaluation et la reconnaissance de l'autorialité. Et renforce l'idée que le « crédit » scientifique de l'auteur ou du contributeur est constitué par la reconnaissance par les pairs de son autorité épistémique ET de sa légitimité scientifique (Broudoux, 2017).

Les revues sont devenues le passage obligé pour la diffusion des connaissances et la mise en place de procédures de validation des productions savantes s'est renforcée à partir du XVIIIe siècle, avec la constitution de comités éditoriaux dans les revues (Bégault, 2007). Puis, les conditions de la validité scientifique ont été transformées par le passage d'un contexte de rareté de l'information scientifique à une ère de profusion à même de renouveler les modes d'autorité académiques (Jensen, 2006). Les pratiques d'évaluation dans le monde pré-web sont régies par des coûts de publication qui ont motivé le passage par des intermédiaires (éditeurs, pairs, etc..). Ces pratiques sont désormais interrogées par l'arrivée de nouvelles métriques bibliométriques et l'importance accordée à de nouveaux acteurs dans le processus évaluatif (Pontille et Torny, 2017).

Au nom de l'open science, les incitations à l'ouverture se relaient dans le monde de la communication scientifique. De l'accès libre aux résultats des recherches à la gestion de sa présence sur les réseaux sociaux, une vague d'initiatives - qui paraissent à leur début subversives - a été absorbée par la plateforme du web2.0 qui vise à agréger des profils en simulant des espaces collaboratifs. Dans ce contexte, nous nous demandons si l'*open peer-reviewing* (OPR)[2] permettra une meilleure reconnaissance de la nature collaborative du travail des chercheurs ou bien s'il ne représente que l'effacement du dernier rempart préservant leur relative autonomie. Car les incitations à la « transparence » révèlent aussi d'autres enjeux liés à la médiatisation : par ex., le référencement web est silencieusement entré dans les pratiques de sélection de recrutement des chercheurs ou de financement des projets.

Afin de mieux connaître ce champ de la sélection des articles, nous proposons de passer par un recensement d'études quantitatives et qualitatives qui décrivent finement les différentes facettes de l'examen par les pairs, leurs variantes et les différences et continuités que son ouverture provoque. Enfin, l'offre des plateformes, celles des mégarevues et des acteurs de la promotion de l'open peer-reviewing est parcourue.

L'ÉVALUATION PAR LES PAIRS

Processus et procédures de l'examen par les pairs

L'évaluation par les pairs (*peer-reviewing*) représente un processus d'examen sélectif d'article ou de communication qui se déroule généralement en amont de la publication et qui se différencie fortement selon les usages disciplinaires. Il est organisé par les éditeurs-chercheurs qui choisissent les membres des Comités de lecture en fonction de numéros thématiques de revues ou d'actes de conférences. Ce processus est marqué par des procédures d'anonymisation des identités des relecteurs-évaluateurs qui visent à maintenir une égalité de traitement entre les auteurs. Cependant, un déséquilibre se produit lorsque les auteurs sont maintenus dans l'ignorance de l'identité des évaluateurs alors que ceux-ci sont à même de les (re)connaître.

Bien que coexistent des variations importantes entre les disciplines, selon les revues et les partenaires, trois modèles de l'évaluation par les pairs sont reconnaissables avant la publication des articles : la connaissance des identités par les auteurs et les évaluateurs (évaluation ouverte), le masquage de l'identité des auteurs pour les évaluateurs (simple-aveugle) auquel s'ajoute celui des évaluateurs entre eux (double-aveugle).

S'il faut garder en mémoire le fait que l'évaluation-tri entre comme paramètre dans l'économie des revues, avec des taux de rejets des articles indicateurs de leur attractivité, les processus d'anonymisation/identification de l'auteur et de l'évaluateur évoluent suivant les contextes éditoriaux et avec le temps. L'augmentation de la production scientifique est corrélée avec les phases de l'automatisation : le papier carbone de la machine à écrire a facilité la circulation des manuscrits au sein des comités de lecture dès la fin du XIXe, avant que les photocopies du XXe ne prennent le relais. A cette période, les revues avaient déjà vécu un basculement : alors qu'elles cherchaient à remplir leurs colonnes, elles ont dû adopter des méthodes pour mieux trier les articles face à l'augmentation de l'offre (Spier, 2002).

La mise en évidence des biais liés à l'évaluation

Le processus d'évaluation par les pairs fait l'objet de critiques, car plusieurs biais remettant en question son intégrité ont été depuis longtemps mis en évidence et l'impartialité - objectif premier des méthodes du double et du simple aveugle - largement remise en cause. En 1977, une étude expérimentale relevait l'absence d'exploration approfondie du « biais de confirmation » qui est la tendance à amplifier et croire des expériences qui soutiennent nos opinions et à ignorer ou discréditer celles qui remettent en cause nos certitudes (Mahoney, 1977). En 2006, une étude portant sur l'évaluation de 1983 posters réalisés pendant trois conférences annuelles avait pour but de tester si les posters soumis par les évaluateurs avaient tendance à recevoir des notes plus élevées que les posters soumis par des non-évaluateurs[3]. Une différence qualitative dans l'évaluation était révélée lorsque les relecteurs étaient eux-mêmes auteurs de posters (Blackburn & Hakel, 2006). En 2012, Carole Lee et al. entreprenaient

une étude récapitulant les caractéristiques du peer-reviewing et ses biais : caractéristiques de l'auteur (prestige, affiliation, nationalité, langue, genre), du reviewer (sur les mêmes soumissions, les relecteurs se divisent en deux classes « stricte » et « clémente ») ; biais liés au contenu, au conservatisme (intolérance à l'innovation ou aux idées neuves), biais de confirmation, contre l'interdisciplinarité, de publication (tendance à publier les résultats positifs) (Lee et al., 2013). Campanario (1998) et Shatz (2004) ont également établi un examen complet des études et des perspectives sur les biais dans l'évaluation par les pairs.

Le débat sur la qualité et l'intégrité des publications scientifiques est récurrent avec la médiatisation de fraudes, de plagiats, la manipulation de données auxquelles s'ajoutent des pratiques questionnables de recherche. Dans ce contexte, les choix d'autorégulation des chercheurs, dont l'examen par les pairs fait partie, sont régulièrement remis en cause. Cependant, les critères d'évaluation sont à distinguer des procédures. Remettre en cause l'évaluation sans se préoccuper de la définition claire des critères est contre-productif.

Les biais de l'évaluation pourraient-ils être résolus par la publicisation des rapports d'examen ? En effet, si des conséquences positives ont été remarquées comme l'amélioration de la qualité des évaluations (Lee et al., 2013), et que l'innovation dans l'évaluation des articles et la clarté du processus de *peer-reviewing* ont été le levier pour la croissance de la mégarevue *Plos One* (Verlaet & Dillaerts, 2017), on peut se demander si des biais susceptibles de s'ajouter aux précédents ne sont pas reconstruits. Ainsi, le biais d'ordonnancement qui est la probabilité plus faible de sélectionner des items classés en bas de liste, serait favorisé par les logiciels de gestion de congrès de type *confmaster.net*. Dans ce système, chaque membre du comité de programme sélectionne les articles qu'il se propose d'évaluer en fonction de son expertise. Or, une étude, prenant en compte la temporalité, conduite par Guillaume Cabanac (2016) a montré que les articles soumis en dernier avaient moins de chance d'être sélectionnés que ceux soumis les premiers.

Les critiques faites à l'évaluation anonymisée

Les procédures d'anonymisation des textes sont rendues partiellement inefficaces aujourd'hui par la médiatisation et l'activité de recherche d'informations sur le web ; les conduites d'auto-référencement par les publiants eux-mêmes — sur les réseaux sociaux numériques tels par exemple, *ResearchGate* et *AcademiaEdu* — accentuent ce phénomène.

Ross-Hellauer (2017) répertorie six critiques récurrentes faites à l'évaluation réalisée avec des identités masquées et des rapports non publiés^[4], depuis la fin des années 1990. A l'incapacité de détecter des erreurs méthodologiques (pour les revues classées par facteur d'impact, un nombre croissant de rétractations en découle), s'ajoute l'invérifiabilité de la « boîte noire » du double aveugle. Des risques de subversion en découlent comme la sélection par les éditeurs d'évaluateurs rejetant les mêmes théories. Ceci peut déboucher sur des manipulations (appropriations d'idées, conflits

d'intérêts non dévoilés). En amont, les biais sociaux sont susceptibles de fausser les évaluations, en termes de genre, nationalité, affiliation institutionnelle, langue, discipline, et en aval, un manque de reconnaissance existe pour le travail invisible effectué par les évaluateurs. Enfin, les resoumissions d'articles reproposés entraînent des pertes de temps dues aux délais d'attente ; de même pour les remarques secrètes des évaluateurs qui ne profitent pas aux chercheurs débutants, ce qui cause des effets de gaspillage.

Il faut ajouter ici une tendance dans les revues qui constatent une augmentation des refus d'évaluer, provoquée par la masse croissante d'articles et la surcharge de travail des spécialistes (Bakker & Traniello, 2019). D'une manière générale, un équilibre est à trouver entre le nombre d'articles soumis et le nombre d'évaluations acceptées pour que le système fonctionne. Une étude réalisée par *Publons* (2018) - un nouvel acteur^[5] qui recense les évaluations faites par les chercheurs et agrège les données qui leurs sont liées - révèle que le ratio nombre d'évaluations par soumission est de 1,95 dans les régions où la science est bien structurée et établie, et nettement plus faible dans les régions émergentes comme la Chine, l'Inde et la Turquie (en moyenne 0,66 avis par soumission). Ce qui révèle de fortes disparités.

L'ÉVALUATION OUVERTE - LES DIMENSION DE L'OPR

Il est commun de faire remonter le démarrage du mouvement d'ouverture du peer reviewing à la fin des années 1990 au *British Medical Journal* (BMJ) lorsqu'une expérimentation menée au sein de cette revue montra qu'il n'existait pas de différence de qualité notable dans les évaluations lorsque les examinateurs s'étaient identifiés (Wolfram, Wang, Park, 2019). Pour Smith (1999), l'objectif était de renverser le processus de sélection en simple-aveugle : puisque les évaluateurs connaissaient l'identité des auteurs, pourquoi les auteurs ne connaîtraient-ils pas l'identité de leurs examinateurs ? S'appuyant sur une série d'essais contrôlés randomisés, la revue BMJ décide de publier le nom des évaluateurs au même titre que ceux du cadre éditorial ; l'objectif suivant étant de publier le nom des examinateurs à la fin des articles. Citant le world wide web naissant, l'auteur imagine l'OPR tel qu'il est aujourd'hui matérialisé par des systèmes gérant le workflow éditorial ; un processus qui gagnerait en scientificité, observable de l'extérieur, valorisant le travail des évaluateurs et dans lequel lecteurs et auteurs seraient reliés plus étroitement. Cependant, dans le modèle initial de Smith, le contenu des évaluations reste confiné aux pairs.

Une grande variété de modèles

L'évaluation ouverte (OPR) possède une multiplicité de formes dont il est difficile d'extraire une définition unifiante. Dans une étude récente sur l'adoption de l'OPR dans les revues en accès libre, Dieter Wolfram et al. (2019) distinguent celles fondées sur la connaissance des identités de celles fondées sur la publication des rapports. Le modèle de l'OPC (*open peer commentary*) peut conserver une évaluation initiale basée sur un anonymat plus ou moins restrictif pour ensuite ouvrir des discussions entre auteurs,

évaluateurs et public (Verlaet et Chante, 2017). Didier Torny (2018) identifie l'évaluation des articles de revues qui se déroule après leur publication (*post publication peer review* ou PPPR). Cette distinction révèle cependant un entre-deux important, constitué par la publication des rapports alors que les articles sont encore en cours d'évaluation.

Les usages vont de la simple divulgation des identités, en passant par la communication des évaluations dans des systèmes où les rapports des évaluateurs sont publiés en rapport avec les articles, jusqu'à la publication ouverte des rapports où évaluateurs et auteurs sont susceptibles de s'entre-commenter. L'ouverture maximum est celle qui permet à d'autres commentateurs que les « experts invités » de contribuer. C'est le modèle « à deux étapes », en usage à la revue ACP[6] où les articles proposés sont discutés dans un forum ouvert à tous puis sélectionnés par la revue à partir de rapports anonymes ou identifiés. Pöschl (2012) qui a analysé ce modèle dit qu'il intègre la « rigueur du modèle traditionnel » et les « vertus de la transparence et de l'auto-régulation ». Parmi les questions-clés qu'il estime nécessaire de prendre en compte dans la conception de systèmes ouverts d'évaluation en sciences, il note que l'évaluation ouverte en pré-publication est préférable à celle de post-publication, car ni l'auteur ni l'évaluateur ne peuvent apporter des améliorations à un article déjà publié, les efforts de l'évaluateur pour proposer des ajouts, des corrections ou des références étant rendus inutiles.

Les autres points sur lesquels Pöschl insiste visent à rassurer les auteurs et les évaluateurs : seuls les rapports et conseils concernant les manuscrits acceptés pour publication devraient être rendus publics, afin de donner une chance aux auteurs de resoumettre leurs travaux ; l'option de l'anonymat doit être préservée pour les évaluateurs ayant peur d'apparaître ignorants ou irrespectueux dans leurs critiques bien qu'un « système optimal » devrait encourager les évaluations signées et authentifiées. La question de l'évaluation en tant que processus continu d'amélioration d'un article, mis à jour au fur et à mesure de ses améliorations, est aussi posée, afin de travailler en profondeur les articles les plus méritants. Cependant, les auteurs aimant mettre un point final à leurs articles, cette possibilité apparaît peu exploitable.

On notera que Pöschl (2012) brandit la transparence comme une idéologie bienfaisante au même titre que la « liberté d'expression » dans leur capacité à maintenir et renforcer l'objectivité scientifique. Ce serait même l'unique « moyen durable » de rechercher « l'objectivité » dans un équilibre des pouvoirs et des intérêts.

Les types d'évaluation ouverte

Sept manifestations de l'évaluation ouverte classés en trois types ont été repertoriés par Tony Ross-Hellauer (2017) avec une revue de la littérature sur un corpus de 122 définitions d'*open peer-review*.

La **connaissance des identités** (*Open identities*) pour chacun des protagonistes (auteurs et évaluateurs) n'implique cependant pas la divulgation des évaluations et

commentaires à un plus large public.

La **publication des évaluations** possède quatre directions dans l'ouverture de la participation mais n'implique pas systématiquement l'identification des évaluateurs. Les rapports (*Open reports*) sont publiés en regard de l'article évalué. L'interaction ouverte (*Open interaction*) est autorisée et encouragée entre auteurs et évaluateurs mais aussi entre les évaluateurs. La participation ouverte (*Open participation*) est la communauté élargie pouvant participer à l'ensemble du processus, elle concerne donc aussi les lecteurs « ordinaires ». Mais est-ce encore du « *peer reviewing* » puisque le processus ne se déroule plus entre pairs ? La dernière version ouverte à commentaires (*Open final-version commenting*) concerne la version finale enregistrée de l'article.

La **portabilité** est l'autonomisation du processus d'évaluation en un service détaché de l'acte de publication. Les auteurs publiant des « manuscrits » dans des dépôts d'archives ouvertes, avant ou pendant leur évaluation rentrent dans cette catégorie (*Open pre-review manuscripts*). Lorsque l'évaluation est découplée, c'est-à-dire réalisée par une entité organisationnelle différente de celle du lieu de publication, Ross-Hellauer la nomme plateforme ouverte (*Open platform*).

Relevons ici que la plateforme entre comme acteur sociotechnique de l'évaluation. Elle est utilisée comme argument implicite de transparence pour convaincre car toutes les actions qui s'y déroulent sont enregistrées et vérifiables.

Pour Schmidt et al. (2018), l'évaluation ouverte installerait « une assurance de qualité constructive » reposant sur la confiance mutuelle, le respect et l'ouverture à la critique. Elle accroîtrait le sens des responsabilités en facilitant une discussion plus large et inclusive. Enfin, elle apporterait une reconnaissance aux évaluateurs avec la citation possible de leurs remarques.

Comme le facteur temps est un reproche fait à l'examen par les pairs traditionnels (les Allers-Retours entre évaluateurs et auteurs prennent du temps), on peut se demander ce qu'il en est avec l'OPR. Certaines recherches suggèrent qu'elle serait susceptible d'augmenter le temps nécessaire à l'examen des articles (Tennant et al., 2017), les répondants souhaitant apporter plus de qualité rédactionnelle à des rapports voués à être rendus public. D'autres indiquent le contraire, comme l'étude de Bravo et al. (2019) sur 9 220 soumissions et 18 525 évaluations pour cinq revues entre 2010 et 2017 : au fur et à mesure que le nombre d'observations écrites augmentait, le temps de rotation entre évaluations restait stable. Même en cas d'interactions entre évaluateurs, les différences sont restées minimales.

Dans leur guide pour la mise en place d'OPR, Ross-Hellauer & Görögh (2019) incitent à la création d'identifiants (ORCID) et à la connexion avec Publons, un service qui crédite les évaluateurs de points, dans une échelle de la visibilité et de la participation pouvant aller jusqu'à obtenir des prix. Cependant, les auteurs admettent que l'OPR est souvent mal comprise et les enquêtes sur les attitudes des chercheurs révèlent d'importants obstacles à sa mise en œuvre.

Problèmes de la publication des rapports, de l'anonymat et de l'identification des rapporteurs

L'OPR se divise en deux grandes tendances : la première consiste à rendre publics les textes des évaluations tout en gardant l'anonymat des contributeurs (auteurs et relecteurs), la seconde consiste à publier les noms des évaluateurs en regard des évaluations. Pour détecter les tendances d'acceptation ou de rejet actuel de l'ouverture de l'évaluation par les pairs, il est difficile de percevoir une évolution des comportements et des usages sur une période de temps aussi courte qu'une vingtaine d'années, comme le montrent les résultats des études suivantes.

Dans sa thèse, Nathalie Pignard-Cheynel (2004) faisait état d'une enquête menée par l'ALPSP[7] en 2000 révélant que 88% des répondants préféraient dissimuler l'identité des évaluateurs lors de l'envoi des rapports d'expertise aux auteurs, dans l'objectif de protéger l'évaluateur d'éventuelles plaintes ou réclamations de la part de l'auteur. Une récente étude pilote portant sur les effets de l'OPR sur le comportement des évaluateurs, à partir de cinq revues d'*Elsevier*, confirme cette tendance. Bravo et al (2019) remarquent que seuls 8,1% des évaluateurs ont accepté de révéler leur identité dans les rapports publiés. D'après ces auteurs, ces résultats suggèrent qu'un examen par les pairs ouvert ne compromet pas le processus, du moins lorsque les arbitres sont en mesure de protéger leur anonymat.

Une étude massive et mondiale multidisciplinaire (3040 réponses sur 39000 e-mails) portant sur l'évolution de l'examen par les pairs a présenté un paysage divisé auprès d'auteurs-évaluateurs-éditeurs (Ware, 2008). Questionnés sur son évolution récente et ses modifications comme la croissance de l'examen en double-aveugle, l'introduction de l'évaluation ouverte ou celle qui se déroule après que les articles soient publiés, les répondants ont préféré le double-aveugle pour plus de la moitié. À la question de savoir laquelle parmi les quatre types d'évaluations était leur option privilégiée, ils ont préféré l'examen en double aveugle (56%), suivies de 25% pour le simple-aveugle, de 13% pour les évaluations ouvertes et de 5% pour les examens postérieurs à la publication. L'OPR agissait comme un repoussoir pour de nombreux évaluateurs, 47% d'entre eux affirmant que la divulgation de leur nom à l'auteur les rendrait moins enclins à s'engager dans une évaluation.

Cette étude confirme les enquêtes pré-existantes de Van Rooyen et al. (1998) qui constataient que le démasquage des relecteurs augmentait les chances de ceux-ci de décliner le travail d'évaluation (23% contre 35%) alors que la plupart des auteurs préféreraient connaître l'identité des relecteurs (55% en faveur contre 26%) ou celle de Melero et López-Santoveña (2001) qui remarquaient que 75% des évaluateurs étaient pour la préservation de leur anonymat (Nobarany & Booth, 2017). Ces derniers auteurs ont réalisé une étude comparative entre les politiques d'anonymat pratiquées par les éditeurs (essentiellement STM) et leur utilisation de l'automatisation sur les plateformes dédiées. Leur constat est que les politiques d'anonymisation peuvent ressortir renforcées mais tout aussi bien être réduites ; une souplesse d'actions est donc nécessaire pour que chaque système automatisé de publication puisse s'adapter aux

exigences disciplinaires et aux choix éditoriaux.

Un examen des études d'*Elsevier* sur plus de 3000 chercheurs actifs répondants, conduit par Mulligan & van Rossum (2014), a montré que l'anonymat comptait beaucoup pour les chercheurs pendant le processus d'évaluation. Les questions portaient sur des options d'évaluation allant du double aveugle au simple aveugle en passant par différentes formes d'évaluation ouverte entre pairs. Les auteurs ont découvert qu'au fur et à mesure que la transparence augmentait, la probabilité que les auteurs soumettent des articles diminuait. 46% ont déclaré qu'il était peu probable qu'ils soumettent un article à une revue dans laquelle le nom de l'évaluateur était publié en parallèle des rapports (identité ouverte). Outre le fait que les auteurs n'aimaient pas l'idée que leurs erreurs puissent être rendues publiques, une proportion de 49% a déclaré qu'elle n'effectuerait pas d'évaluation dans une revue qui proposerait un OPR.

Cette réticence se confirme par une étude pluridisciplinaire récente menée par Ross-Hellauer, Deppe & Schmidt (2017) à partir du workshop du projet européen OpenAire2020 intitulé « *Open Peer Review - models benefits and limitations* » qui s'est tenu en juin 2016 à Göttingen (Allemagne). Sur les trois quarts des répondants (76,2%) ayant déjà pris part à des OPR, 50,8% exprimaient une opinion défavorable à l'idée de publier le nom des évaluateurs en face des rapports.

L'étude des commentaires

Wakeling et al. (2018) ont présenté une analyse quantitative et qualitative des commentaires sur les articles publiés par le mégajournal, *Plos One*, entre 2003 et 2016. Ils proposent une typologie des commentaires divisée en deux catégories non exclusives : une partie procédurale qui concerne les commentaires non liés au contenu du document, mais plutôt faisant référence au processus de publication, à l'expression, à la composition, aux références ; une partie académique qui concerne les commentaires s'intéressant d'une manière ou d'une autre au contenu scientifique de l'article. Leur constat est que la moyenne du taux de commentaires déjà faible était en diminution depuis 2010 ; le pourcentage d'articles avec commentaire est de l'ordre de 7,4 % avec en moyenne 1,9 commentaire par article. La typologie des commentaires développée pour cette recherche (tableau traduit en Annexe 1) révèle que près d'un tiers des commentaires (29,1%) avaient été formulés par un éditeur et environ un quart seulement (23,8%) par l'auteur de l'article. Les commentaires produits par les lecteurs étant les plus fréquents (38,3%), mais une variation importante existe entre les différents titres de Plos. Environ la moitié des commentaires contient une discussion théorique de l'article mais ces discussions se concentrent essentiellement sur la validité technique du document.

L'interprétation des commentaires par Wakeling et al., révèle que la catégorie « Discussion » est loin de rejoindre le cœur revendiqué par l'OAMJ de la validité scientifique et du techniquement solide en méthodologie et en analyse. Les commentaires tendaient à focaliser sur des questions aux auteurs, les tendances du

futur travail, ou des débats autour des définitions ou de la terminologie. Même si ces discussions sont intéressantes, elles ne représentent pas des critères sur lesquels prendre des décisions d'acceptation ou de rejet (Wakeling et al., 2018). Au final, cette étude révèle que les commentaires sont surtout réalisés par les lecteurs externes au processus de peer-reviewing. Alors que la rhétorique des éditeurs de l'OAMJ est basée sur le fait que « la communauté décide » (Spezi et al., 2018) des articles importants, il semblerait pour Wakeling et al. qu'en pratique il n'existe aucun moyen de regrouper et de communiquer efficacement les points de vue de la « communauté ».

Le résultat s'apparente donc davantage à « le chercheur décide », car il incombe à chaque lecteur-chercheur de découvrir et d'évaluer des articles. Alors que les taux de soumission globaux aux mégarevues pourraient atteindre un plateau, la vision des mégarevues de démocratisation de la science repose sur le fait que ces articles sont non seulement publiés, mais également lus et suivis publiquement. Il semble donc essentiel que les éditeurs de l'OAMJ continuent d'étudier de nouveaux outils d'évaluation et de découverte d'articles.

En 1998, Bingham et al. remarquaient que les critiques post-publication - souvent brèves et concernant des points de détail - effectuées par les lecteurs en ligne pouvaient apporter de précieux commentaires, mais ne suffisaient pas à remplacer une évaluation complète par des pairs. Vingt ans plus tard, il ne semble pas que cette considération ait notablement évolué. S. Nobarany & K. S Booth (2017) notaient ainsi qu'en 2006, la revue *Nature* obtenait des résultats similaires et que ce type d'examen suscitait peu d'intérêt de la part des auteurs ou des relecteurs. Cette période se caractérise d'ailleurs par l'abandon des blogs scientifiques par les grands éditeurs comme *Nature* (Broudoux, 2013).

Lors d'un examen par les pairs traditionnels, on considère la convergence des évaluations des examinateurs indépendants comme positive, c'est un signe de fiabilité de l'examen des propositions. Cependant, l'étude sur l'open peer review réalisée par Bornmann et Daniel (2010) qui ont exploré ce principe de fiabilité inter-évaluateurs indique que les évaluations réalisées par commentaires ouverts de la revue ACP possédaient un niveau faible de convergence et comparable à celui des processus traditionnels d'évaluation par les pairs. Un haut niveau de convergence ne serait d'ailleurs pas toujours souhaitable, puisqu'il pourrait aussi révéler une redondance entre évaluateurs qui n'auraient pas été sélectionnés pour leur diversité (Bailar, 1991) cité par Bornmann et Daniel (2010). En résumé, les auteurs estiment qu'un degré faible de fiabilité inter-évaluateurs ne doit pas être envisagé comme une indication de la qualité du processus d'examen par les pairs, mais comme une de ses caractéristiques générales pouvant éventuellement avoir un effet positif sur la validité prédictive des décisions de sélection des manuscrits par les éditeurs. De ce point de vue, on peut donc affirmer que la non convergence de l'évaluation ouverte n'est pas un signe de manque de fiabilité.

Par ailleurs, il faut relever que les types de recommandations des évaluateurs varient selon le degré d'ouverture des évaluations et de leur anonymat. Lorsque leur nom est

publié, le pourcentage de rejet est faible et les critiques mineures plus nombreuses que les critiques majeures (Bravo et al., 2019). Les auteurs remarquent ainsi que ceux qui écrivent les critiques plus positives seront aussi les plus enclins à révéler leur identité plus tard en tant que signal de réputation auprès des auteurs et de la communauté.

TPOLOGIE DES ACTEURS DE L'OPR

L'intégration par les mégarevues d'un workflow éditorial modélisant l'évaluation

La fondation par les grands éditeurs de méga-revues multidisciplinaires est le signe d'une conversion de leur modèle éditorial au *Gold Open Access* (CNRS, 2015). La Dist du CNRS (2015) révèle ainsi que près de 80% des nouveaux titres lancés sur le marché sont des revues en *Open Access*. Ce modèle s'impose comme réponse optimale aux flux croissants de la publication scientifique car les éditeurs renvoient aux mégarevues les propositions qu'ils refusent, les articles ne quittant pas ainsi le giron de l'éditeur (Boukacem, 2014).

L'évolution des critères d'évaluation par les pairs est rendue visible par la mégarevue *Plos One*, qui prend en compte des résultats négatifs « non recevables » par les revues classiques (Verlaet & Dillaerts, 2017). De nouvelles stratégies de publication couplées à de nouveaux modèles d'évaluation par les pairs associés sont devenues possibles : la barrière à la sélection étant peu élevée, un grand nombre d'articles est publié. Les auteurs y trouvent une perspective plus large et plus rapide de publication de leurs articles et les éditeurs gagnent des revenus proportionnels au nombre d'articles publiés (Boukacem, 2014). Il est demandé aux examinateurs de *Plos* de fonder leurs recommandations d'acceptation ou de rejet sur le bien-fondé et la robustesse de la recherche, notamment la méthodologie et la validité des résultats. Pour *Plos*, la fonction de publication des « reviews » est proposée en option à l'auteur et à l'évaluateur pour la signature de ses évaluations, le processus d'évaluation étant par défaut en « simple aveugle ». Lorsqu'un article est accepté, l'auteur est alors invité à publier l'historique de ses évaluations qui comprend la lettre de décision éditoriale complète pour chaque révision avec ses évaluations ainsi que les réponses aux commentaires des évaluateurs, y compris les pièces jointes.

Dans une méga-revue, même si les orientations éditoriales subsistent au sein des revues disciplinaires, la qualité seule de l'article soumis devrait être prise en compte (Guédon, 2014). Selon la philosophie de *Plos* les examinateurs ne devraient pas juger de la nouveauté, de la pertinence ou de la signification de la recherche, évaluation qui devrait être laissée au lecteur et à la communauté au sens large. En se concentrant sur la rigueur et la solidité (bien-fondé) de la recherche, la mégarevue vise à garantir la publication de tous les résultats utiles et à empêcher toute évaluation subjective sur la signification ou la pertinence d'une étude (Horbach, Halffman, 2018).

Mais cette technicité objective du bien-fondé est remise en question par des résultats

qui suggèrent que d'autres critères pourraient influencer les décisions des relecteurs, engendrant des déviations au modèle OAMJ (*Open-Access Mega-Journals*) (Spezi et al., 2018). En particulier, l'étude par entretiens de ces auteurs a révélé que les exigences des éditeurs pour qu'un article soit « digne » d'être publié entraînaient des jugements d'importance et de nouveauté ; ils ont relevé également que les évaluateurs et rédacteurs mettaient en pratique des critères traditionnels d'évaluation pour justifier de la publication ou du rejet des articles, comme l'originalité et la pertinence.

Les modèles des revues et plateformes de publication centrées sur l'évaluation

Dans ces modèles, l'activité de la revue ou de la plateforme est concentrée sur l'évaluation, encodée par un workflow éditorial. Elle intègre un écosystème visant différents objectifs pouvant aller jusqu'au référencement et la recommandation. Deux modèles sont identifiables, celui de la revue ou mégarevue pouvant s'instituer en plateforme (comme F1000Res) et celui du découplage où l'étape d'évaluation est déconnectée de la publication dans une revue (comme Peer Community In).



Tableau 1 - Les modèles de revues et des plateformes face à l'OPR

Le modèle collaboratif basé sur le libre-accès gérant l'ensemble du processus de publication

Acquise par le Groupe *Taylor & Francis* en janvier 2020, la plateforme *F1000Research* a été pionnière dans l'OPP. Elle fait partie de la galaxie F1000 et couvrirait aujourd'hui 40 disciplines et environ 3 500 revues. Le choix de la Commission Européenne, en mars 2020, de F1000 pour concevoir l'*European Commission Open Research Publishing Platform* destinée à évaluer, publier et gérer l'archivage des articles financés par Horizon 2020[8] révèle des incitations qui pourraient devenir des obligations.

Les auteurs publient directement leurs articles qui sont immédiatement lisibles et citables. La version 2020 explicite le workflow éditorial[9] : les auteurs et l'équipe éditoriale assistée par un algorithme (*Reviewer Finder Tool*) proposent des examinateurs experts qui sont in fine sélectionnés par les auteurs. Les rapports des évaluateurs identifiés sont publiés à côté de l'article, avec les réponses des auteurs et les commentaires des utilisateurs enregistrés. Les auteurs sont encouragés à publier les versions révisées de leurs articles. *F1000Research* génère des citations typographiquement normées avec une granularité fine : toutes les versions d'un article sont liées et citables, de même que les commentaires réalisés par les évaluateurs.

C'est aussi l'objectif des OAMJ comme *Frontiers* de favoriser les interactions entre auteurs et relecteurs afin d'arriver à un consensus sur les moyens d'améliorer la qualité des articles proposés. Dans ce modèle à étapes, la sélection est séparée de la discussion. A l'étape de pré-sélection, les identités des évaluateurs restent secrètes.

Puis, les rapports d'évaluation sont publiés dans le forum, ce qui ouvre l'étape de discussion entre auteurs et évaluateurs, sous la houlette des éditeurs. Ce n'est qu'à la publication finale de l'article que les identités des évaluateurs seront rendues visibles à tous. Dans ce modèle, quatre rôles d'éditeurs encadrent la totalité du processus et jouent si besoin un rôle d'arbitres entre évaluateurs et auteurs sur le forum.

Le modèle de découplage scindant l'activité d'évaluation de celle de publication

Il se décline en pré-publication avec notamment les épi-revues construites sur les serveurs d'archives ouvertes hébergeant les pré-print et en post-publication avec les articles publiés ouverts à commentaires.

La plateforme gratuite et non commerciale de recommandation de pré-prints *Peer Community In* (PCI[10]) ne vise pas le suivi des activités des évaluateurs, mais contribue à valoriser leur activité. Le système s'appose à un serveur de preprint et lorsque les propositions sont validées, les évaluations sont publiées et servent de recommandation pour leur publication dans une revue ou un serveur d'archives ouvertes. Il s'agit d'un système de publication alternatif basé sur la recommandation des preprints et leur évaluation, le tout dans une approche disciplinaire car les communautés de pairs s'agrègent par spécialité : Peer Community in Evolutionary Biology (PCI Evol Biol), Peer Community in Ecology (PCI Ecol), Peer Community in Computational Statistics (PCI Comp Stat), et Peer Community in Paleontology (PCI Paleo).

PRElights est un modèle collaboratif fondé par The Company of Biologists – société savante créée en 1925 (UK) et éditrice de 5 revues. Les évaluateurs - invités à demander eux-mêmes leur intégration – sont chargés de sélectionner les pre-print pour les intégrer à une interface d'examen, où ils seront susceptibles de devenir l'objet d'une conversation globale et non médiatisée (« ouvert » est ici considéré comme public + non anonyme + scriptible pour tous les chercheurs authentifiés).

En post-print, la plateforme multidisciplinaire PubPeer ouvre à tous la possibilité de commenter des articles déjà publiés, en laissant à chacun le choix de l'identification ou de l'anonymat. Récemment en 2020, c'est par *PubPeer* que les faiblesses d'un article cosigné par le Pr. Raoult et publié par *The Lancet* ont été portés à la connaissance du grand public, article ensuite retiré de cette revue. *PubPeer* permet de questionner les articles et donne la possibilité de faire connaître des fraudes, mais autorise aussi le lancement de campagnes de dénonciation (voir la controverse concernant Catherine Jessus[11]). L'anonymat des commentateurs relance alors le débat sur le maintien de l'objectif de qualité du processus d'évaluation. D. Torny (2018) qui a analysé cette plateforme s'interrogeait sur ses fonctions : science vigilante, club de revue ou lancement d'alertes ?

Le modèle de découplage basé sur les auteurs

Ce modèle révèle une orientation qui centre l'activité d'évaluation autour de l'auteur.

Researchers.one[\[12\]](#) fondée par les jeunes chercheurs Harry Crane et Ryan Martin (2018) est une plateforme multidisciplinaire de preprint d'auteurs, ouverte à commentaires. Est mise en valeur l'autonomie des chercheurs à poursuivre leur passion, leur autorité pour développer et disséminer eux-mêmes leur travail et l'accès pour s'engager dans la « communauté internationale des chercheurs ». Dans ce modèle, le fait de déposer un article le rend accessible l'article à la lecture et au commentaire par tous.

The Winnover[\[13\]](#) est emblématique du devenir économique des initiatives du libre-accès. La plateforme d'écriture et édition d'articles scientifiques Authorea l'acquiert en 2016 avant d'être rachetée à son tour par *Wiley*. Après le dépôt par l'auteur d'un article ou de tout autre écrit qu'il estimait légitime de rendre public, celui-ci était automatiquement publié et ouvert à « évaluation » aux membres enregistrés sur la plateforme. La version finale était stabilisée par l'auteur qui achetait alors un DOI.



Tableau 2 - Déclinaisons OPR de l'évaluation, des modes d'ouverture et des coûts par revue ou plateforme

Le modèle de la plateforme pour évaluateurs

Publons est un service valorisant le travail des évaluateurs, en recensant les rapports d'évaluation et en reliant les articles évalués aux données du *Web of Science* (WoS). Il est possible pour un évaluateur de publier ses rapports et de suivre l'évolution d'un article (par exemple un article refusé qui serait reproposé et accepté par une autre revue). Surveillant ses tableaux de bord, l'évaluateur se voit attribuer des points et des prix. *Publons*, racheté en 2017 par *Clarivate Analytics* (société acquise par *Onex*[\[15\]](#) en 2016), est un service qui s'ajoute à la base de données bibliographiques de données de recherche WoS, à *ScholarOne*, service de gestion de manuscrits (processus de soumission, d'évaluation et de suivi d'articles), ainsi qu'au *Journal of Citation Report* (JCR), qui fournit des informations bibliométriques sur plus de 11 000 revues scientifiques dans plus de 230 disciplines du domaine des sciences et des sciences sociales (Wikipédia, entrée JCR).

CONCLUSION

Discussion

Les différents modèles étudiés nous permettent de préciser le classement proposé par

Wolfram et al. (2019). Le modèle innovant de la plateforme ouverte collaborative est celle qui mise sur l'autoritativité de l'auteur et celle du lecteur, reposant sur le contrat implicite de la solidarité à l'intérieur d'une discipline. Les revues comme *Atmospheric Chemistry and Physics*, les plateformes *F1000Research* ou *Peer Community In* proposent des espaces d'écriture (forums, champs texte dédiés aux commentaires, etc.) qui ont été adoptés par les usages disciplinaires. Elles représentent des tentatives intéressantes de création de nouvelles conditions au déroulement de l'évaluation. Le traitement de l'anonymat y est fortement différencié, laissant le choix aux évaluateurs (opt-in) ou au contraire imposant l'identification comme premier critère.

Deux tendances pour les plateformes d'examen ouvert par les pairs sont visibles : celles qui s'adressent aux auteurs (plateformes de publication de preprint dont l'ouverture à commentaires est susceptible de varier) et celles qui s'adressent aux évaluateurs (plateformes de postpublication ou de découplage).

Cette exploration nous a permis de remarquer la tendance des revues à se modulariser (modèles d'écrits, espaces d'évaluation, de discussion, publication, diffusion, référencement, etc.) devenant ainsi des plateformes avec des services à rentabiliser. On s'interroge aussi sur le modèle économique des revues pratiquant le modèle auteur-payeur (APC) qui en centrant leur activité sur l'évaluation pourrait aussi la transformer en service payant. Dans un article récent, Pierre Carl Langlais et Eprist (2020) ont fait le constat d'une accélération de l'évaluation ouverte à partir de 2017, passant de 247 revues à 618 en 2019. Néanmoins notre article montre que ces chiffres ne révèlent pas une diversité disciplinaire mais plutôt une concentration des revues sur les plateformes de l'OAMJ.

D'une manière générale, l'Open Science dont le fer de lance est l'évaluation ouverte vise à faciliter l'accès aux connaissances pour les entreprises, les administrations, les politiques et les citoyens. Quelles leçons alors à retenir face à la pandémie de l'ouverture des publications ayant trait au coronavirus ? Comment les données visant à fabriquer les vaccins ont-elles été partagées ? Peut-on mesurer ici les effets de la transparence voulue par la Déclaration DORA de 2012 ?

Comme Fitzpatrick (2010) et Verlaet (2017) qui interrogent l'autorité académique, nous notons que dans les modèles explorés, le rôle donné à la communauté reste majeur puisqu'il transforme l'évaluation par les pairs basé sur un processus axé sur le tri-contrôle de ce qui est publié en un processus de filtrage maîtrisé par les pairs auteurs-lecteurs. Dans un tel système, l'investissement des membres dans une discussion ouverte devient le point principal de leur progression et de celui de la communauté dans son ensemble ; ce n'est que de cette manière que l'individu érudit et l'ensemble du domaine réussissent (Fitzpatrick, 2010). L'auteur-créateur-innovateur retrouve une place au cœur d'un dispositif qui pratique un examen non sélectif. Les critères d'importance et de nouveauté s'affaiblissent alors au profit de celui de la qualité (Walker & Rocha da Silva, 2015). Le lien entre auteurs et lecteurs en sortirait renforcé, ce qui contribuerait à l'élargissement du lectorat et à l'enrichissement culturel de la société (Fitzpatrick, 2012). Pontille & Torny (2015) vont plus loin en prétendant que les

lecteurs deviendraient une nouvelle instance de jugement clé pour la diffusion et la validation, renouvelant la définition des pairs et donc le processus d'évaluation. Cependant, ce modèle de la « communauté » pourrait favoriser un phénomène d'endogamie lié à l'effet « petit monde », avec des espaces de polarisation où chacun.e s'inscrirait dans la communauté de ses croyances; de nouveau le biais de confirmation réapparaîtrait.

Compte-tenu de la puissance des acteurs de l'édition scientifique qui promeuvent l'OPR dans un objectif de monétisation des traces de lecture, nous pensons qu'un risque se tient dans le fait que l'activité d'évaluation, actuellement basée sur l'anonymat et le volontariat, pourrait devenir obligatoire et évaluée pour au moins deux raisons qui se renforcent l'une l'autre. La première est la médiatisation provoquée par le référencement de l'évaluateur identifié, la seconde est la course à l'« excellence » par la sélection des « meilleurs »[\[16\]](#). Ce phénomène s'inscrit dans une course à la visibilité, condition à la constitution de capital symbolique spécifique au champ scientifique que P. Bourdieu (1976) avait bien identifié. On peut se demander si la fin de l'anonymat des évaluateurs ne signifie pas alors la fin d'une période qui visait une communication des sciences dépouillée des signes d'auteurs, pour s'incarner à nouveau dans une subjectivation masquée par un appareil d'évaluation exhibé.

Cependant, le modèle autoritatif qui vise à contourner le modèle de filtrage de l'éditeur tirant profit du travail de recherche sans aucun ancrage disciplinaire, uniquement par la gestion automatisée des traces de lecture, représente un espoir de reconnaissance de la nature collaborative du travail des chercheur.se.s. S'ancrant dans les communautés disciplinaires dans lesquelles elles restent confinées, ces initiatives nous montrent qu'une autre voie est possible que celle du « *publish or perish* ». Le rôle de l'éditeur pourrait alors s'en trouver renouvelé. Débarrassé des problèmes d'évaluation, il devrait se concentrer sur son savoir-faire éditorial : créer des collections, redocumentariser en rassemblant des documents selon des critères thématiques, éditorialiser en liant des fragments, en repérant et en mettant en valeur des tendances, des bifurcations, etc.

Les initiatives françaises en sciences humaines et sociales

Les modèles traditionnels d'évaluation et d'autorité demeurent encore la norme dans le champ scientifique. Les bouleversements techniques (numérisation, plateformes, altmetrics, etc.) dégagent de nouvelles possibilités pour la médiatisation des travaux savants. Nous avons identifié dans cet article quelques tendances qui concernent notamment l'OPR dans les sciences.

En France, l'agrégateur de revues, de blogs et éditeur *SHS OpenEdition* a impulsé des initiatives autour de l'évaluation ouverte par les pairs dans des projets européens comme OpenAire, et expérimentant l'ouverture à commentaires avec la revue *VertigO* en 2018. Plus récemment, dans le cadre du projet européen *Hirmeos*, une expérimentation d'évaluation ouverte par les pairs post-publication s'est déroulée en 2019 sur la plateforme *OpenEdition Books*. Cependant ces essais et tentatives restent du domaine de l'expérimentation en SHS, au vu du faible nombre de participants

actifs[17].

Le livrable D3.4 du projet européen *OpenUp* (2018) livre les bonnes pratiques pour réussir un projet d'ouverture de l'évaluation par les pairs. Toute récemment fondée, la *Global Alliance of Open Access Scholarly Communication Platforms* (GLOALL) dont le lancement a été inauguré à l'Unesco en avril 2019 inclura sûrement un volet d'OPR.

Nous constatons que la notion classique de monographie ou même de document est en train de se transformer et de se développer au profit de la notion évolutive de ressource numérique. Au-delà du document numérique (article, monographie) ou de la donnée brute, il s'agit dorénavant de proposer un accès aux protocoles d'expériences (deux nouveaux types d'articles ont été annoncés pour 2021 dans *Plos One* : les Lab Protocols et Study Protocols), aux collections de données (corpus de textes, de sons, d'images et vidéos), et aux logiciels utilisés dans les expérimentations scientifiques. Dans le cas, français, ces questions émergent dans un moment où l'environnement scientifique est en pleine évolution. La mise en œuvre de réformes structurelles dans les universités françaises (HCERES, LPPR, baisse des budgets, individualisation de l'évaluation, rapprochements universitaires, modification du statut des chercheurs, concurrence intra et interuniversitaire) ont modifié les conditions d'accès, d'évaluation et de diffusion des connaissances scientifiques.

La mise en visibilité des processus d'évaluation vient ainsi accompagner un mouvement plus général qui peut aboutir à de nouveaux modes d'argumentation mais aussi à une évolution des modes d'évaluation « scientifique » en SHS. Assistera-t-on alors à un mouvement historique dans les SHS, à savoir une « technicisation » accrue de ses procédures de recherche et l'adoption de pratiques de publication, de fonctionnement de laboratoires, de carrières, plus proches des sciences dures ou expérimentales ?

RÉFÉRENCES BIBLIOGRAPHIQUES

Armbruster, C. (2007). Moving out of Oldenburg's Long Shadow: What is the Future for Society Publishing ? *Learned Publishing*, Vol. 20, No. 4, October 2007. Disponible sur SSRN: <https://ssrn.com/abstract=997819> or <http://dx.doi.org/10.2139/ssrn.997819>

Bakker, T. C., & Traniello, J. F. (2019). Peer-review reciprocity and commitment to manuscript evaluation. Editorial. *Behavioral Ecology and Sociobiology*. 73:38. <https://doi.org/10.1007/s00265-019-2647-2>

Bailar, J.C., (1991). Reliability, fairness, objectivity, and other inappropriate goals in peer review. *Behavioral and Brain Sciences*, 14(1): 137-138.

Bégault B., Usages et pratiques de la publication électronique des résultats de la recherche. Le cas des sciences de l'ingénieur, *Document numérique* 2007/3-4, Volume 10, p. 47-61

Blackburn, J. L., & Hakel, M. D. (2006). An Examination of Sources of Peer-Review Bias.

Psychological Science, 17(5), 378-382.

Bornmann, L., & Daniel, H. D. (2010). Reliability of reviewers ratings when using public peer review: a case study. *Learned Publishing*, 23(2), 124-131.

Boukacem-Zeghmouri, C. (2014). Les couleurs de la publication scientifique. Mutations dans la sous-filière de la revue scientifique STM, analysées par les industries culturelles. *Les Enjeux de l'Information et de la Communication* 15 (1), 49-66.

Bourdieu, P. (1976). Le champ scientifique. In: Actes de la recherche en sciences sociales. Vol. 2, n°2-3, juin 1976. La production de l'idéologie dominante. pp. 88-104. http://www.persee.fr/doc/arss_0335-5322_1976_num_2_2_3454

Bordier, J. (2016). Évaluation ouverte par les pairs: de l'expérimentation à la modélisation. URL : <https://hal.archives-ouvertes.fr/hal-01283582/document>

Bravo, G., Grimaldo, F., López-Iñesta, E., Mehmani, B., & Squazzoni, F. (2019). The effect of publishing peer review reports on referee behavior in five scholarly journals. *Nature communications*, 10(1), 322.

Broudoux, E. (2017). « Autorité épistémique et scientifique à l'épreuve de la mesure des citations » in *Etudes de Communication* n°48, pp. 177-198.

Cabanac, G. (2016). Interroger le texte scientifique. Réseaux sociaux et d'information [cs.SI]. Habilitation à diriger des recherches. Université Toulouse 3 - Paul Sabatier. `

Campanario, J.-M. (1998). Peer review for journals as it stands today, part 2. *Science Communication*. 19 (4): 277-306.

Chaudiron, S., Ihadjadene, M. Existe-t-il un "effet AERES" sur les revues en SIC ? Le cas d'Etudes de Communication IN Fraysse, P., Gardiès, C., Fabre , I. *Sur les sciences de l'information et de la communication : contributions hybrides autour des travaux de Viviane Couzinet*, Editions CEPADUES, pp.68-87, 2017. [hal-01590202]

CNRS-DIST (2015). L'édition de sciences à l'heure numérique : dynamiques en cours, 26p.17

Cotte, D. (2017). Économies scripturaires, formes documentaires et autorité. Réflexions et esquisse d'analyse des architextes de la « science ouverte ». *Communication & langages*, 192,(2), 117-129. doi:10.4074/S033615001701208X.

Ellison, G. (2011). Is Peer Review in Decline? *Economic Inquiry*, 49(3), 635-657. <https://doi.org/10.1111/j.1465-7295.2010.00261.x>

Fitzpatrick, K. (2010). Peer-to-peer Review and the Future of Scholarly Authority. *Social Epistemology*, 24(3), 161-179.

Fitzpatrick, K. (2012). Giving It Away: Sharing and the Future of Scholarly Communication. *Journal of Scholarly Publishing*. 43(4):347. <https://hcommons.org/deposits/item/mia:579/>

Ford, E. (2018). Scholarship as an Open Conversation: Using Open Peer Review in Library Instruction. In *the Library with the Lead Pipe*.

Görögh, E., Ross-Hellauer, T., Schmidt, B., Bardi, A., Casarosa, V., Manghi, P., ... & Woutersen-Windhauer, S. Deliverable D3. 4-Open Peer Review: Good practices and lessons learned. URL : http://openup-h2020.eu/wp-content/uploads/2018/08/OpenUP_D3.4_Open-Peer-Review.-Good-practices-and-lessons-learned.pdf

Granger, S. (2018). Valoriser le travail des reviewers. Retours d'expérience de chercheurs sur le peer review, Publons. Compte-rendu de la journée d'étude du 27 mars 2018 « Le peer review à l'ère de l'open science : enjeux et évolutions » à Bordeaux. <https://urfistinfo.hypotheses.org/3149>

Guédon, J.-C. (2001). A l'ombre d'Oldenburg : bibliothécaires, chercheurs scientifiques, maisons d'édition et le contrôle des publications scientifiques. ARL Meeting, Toronto, mai 2001.

Guédon J.-C. (2014). "Le Libre Accès et la « Grande Conversation » scientifique", in E. Sinatra Michael, Vitali-Rosati Marcello (édité par), *Pratiques de l'édition numérique*, collection « Parcours Numériques », Les Presses de l'Université de Montréal, Montréal, p. 111-126, ISBN: 978-2-7606-3202-8.

Hodonu-Wusu, J. O. (2018). *Open Science: A Review on Open Peer Review Literature*. Library Philosophy & Practice.

Horbach, S. S., & Halffman, W. W. (2018). The changing forms and expectations of peer review. *Research integrity and peer review*, 3(1). URL : <https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-018-0051-5>

Jensen, M. (2006) (En ligne). Authority 2.0 and 3.0: The Collision of Authority and Participation in Scholarly Communications. Disponible sur http://www.michaeljonjensen.com/authority_3_0.pdf

Langlais, P.-L. (2016). Etude critique des nouveaux modes « d'éditorialisation » de revues scientifiques en accès-ouvert. [Rapport de recherche] Bibliothèque Scientifique Numérique. hal-01388556

Langlais, P. -L., Eprist, (2020). « La conversation scientifique ouverte ». *Analyse I/IST* n°29 - Mars 2020. URL : https://www.eprist.fr/wp-content/uploads/2020/03/EPRIST_I-IST_Note-Synthese_Evaluation-ouverte_Mars2020.pdf

Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *JASIST*, 64(1), 2-17. <https://doi.org/10.1002/asi.22784>

Mahoney, M. J. (1977). Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research*, 1(2), 161-175. URL : <https://doi.org/10.1007/BF01173636>

Miège, B. (2015). *Contribution aux avancées de la connaissance en Information-Communication*. INA Editions. 239 p.

Mulligan, A., & van Rossum J. (2014). What you think about the peer-review process: We turn the spotlight on the results of Elsevier studies reflecting the views of more than 3000 active researchers. URL : <http://www.elsevier.com/reviewers-update/story/peer-review/what-you-think-about-the-peer-review-process>.

Nobarany, S., & Booth, K. S. (2017). Understanding and supporting anonymity policies in peer review. *Journal of the Association for Information Science and Technology*, 68(4), 957-971. URL : <http://syavash.nobarany.com/files/taxonomy.pdf>

OpenUP. Deliverable D3.4-Open Peer Review: Good practices and lessons learned. Commission Européenne. URL : http://openup-h2020.eu/wp-content/uploads/2018/08/OpenUP_D3.4_Open-Peer-Review.-Good-practices-and-lessons-learned.pdf

Pignard-Cheynel, N. (2004). La communication des sciences sur Internet : stratégies et pratiques. Thèse de Doctorat en sciences de l'information et de la communication. http://science.societe.free.fr/documents/pdf/These_PIGNARD-6.pdf

Pontille, D., Torny, D. (2014). The Blind Shall See! The Question of Anonymity in Journal Peer Review. *Ada: A Journal of Gender, New Media, and Technology*, University of Oregon Libraries, 2014 URL : <https://adanewmedia.org/2014/04/issue4-pontilletorny/>

Pontille, D., Torny, D. (2015). From manuscript evaluation to article valuation: the changing technologies of journal peer review. *Human Studies*, 38(1), 57-79.

Pontille, D., Torny, D. (2017). Infrastructures de données bibliométriques et marché de l'évaluation scientifique. Pierre-Michel Menger; Simon Page. Big data et traçabilité numérique. Les sciences sociales face à la quantification massive des individus, Collège de France , pp.105-120.

Publons, (2018). Global state of peer review 2018, <https://doi.org/10.14322/publons.GSPR2018>

Polka, J. K., Kiley, R., Konforti, B., Stern, B., & Vale, R. D. (2018). Publish peer reviews.

Ross-Hellauer T. (2017). What is open peer review? A systematic review [version 1;

referees: 1 approved, 3 approved with reservations]. F1000Research 2017, 6:588 (<https://doi.org/10.12688/f1000research.11369.1>)

Ross-Hellauer, T., Deppe, A., & Schmidt, B. (2017). Survey on open peer review: Attitudes and experience amongst editors, authors and reviewers. PloS one, 12(12), e0189311.

Ross-Hellauer, T., Schmidt, B., & Kramer, B. (2018). Are funder Open Access platforms a good idea? SAGE Open, 8(4), 2158244018816717.

Ross-Hellauer, T., & Görögh, E. (2019). Guidelines for open peer review implementation. Research Integrity and Peer Review, 4(1), 4.

Sadeghi, A., Capadisli, S., Wilm, J., Lange, C., & Mayr, P. (2018). Automatically Annotating Articles Towards Opening and Reusing Transparent Peer Reviews. arXiv preprint arXiv:1812.01027

Schmidt, B., Ross-Hellauer, T., van Edig, X., & Moylan, E. C. (2018). Ten considerations for open peer review. F1000Research, 7.

Schöpfel, J., Fabre, R. (2019). La question des revues dans la science ouverte : une approche fonctionnelle. I2D - Information, données & documents, A.D.B.S., 2019, n°2 (2), pp.109. [10.3917/i2d.192.0109](https://doi.org/10.3917/i2d.192.0109). [hal-02403939](https://hal.archives-ouvertes.fr/hal-02403939)

Segado-Boj, F., Martín-Quevedo, J., & Prieto-Gutiérrez, J. J. (2018). Attitudes toward open access, open peer review, and altmetrics among contributors to Spanish scholarly journals. Journal of scholarly publishing, 50(1), 48-70

Shatz, D. (2004). Peer review: A critical inquiry. Lanham, MD: Rowman & Littlefield.

Smith, R. (1999). Opening up BMJ peer review. A beginning that should lead to complete transparency. 318 : 4. <https://doi.org/10.1136/bmj.318.7175.4>.

Spezi, V. Wakeling, S., Pinfield S., Fry, J., Creaser, C., & Willett, P. (2018). « Let the community decide »? The vision and reality of soundness-only peer review in open-accessmega-journals , Journal of Documentation, Vol. 74 Issue: 1, pp.137-161, <https://doi.org/10.1108/JD-06-2017-009>

Spier, R. (2002). The history of the peer-review process. Trends in Biotechnology, 20(8), 357-358.

Tennant, J. et al. (2017). A multi-disciplinary perspective on emergent and future innovations in peer review. F1000Res. 6, 1151.

Torny, D. (2018). Pubpeer: vigilante science, journal club or alarm raiser? The controversies over anonymity in post-publication peer review. PEERE International Conference on Peer Review, PEERE COST Network, Mar 2018, Rome, Italy.

Verlaet, L., Dillaerts, H. (2017). Enjeux des « revues hypermédiatisées » pour l'édition scientifique. dans Kembellec G., Broudoux E. Ecrilecture augmentée dans les communautés scientifiques : humanités numériques et construction des savoirs, ISTE Editions, 2017, p. 145-163 [halshs-01476924](#)

Verlaet, L., et Chante, A. (2017). « La notion d'autorité dans l'édition de la recherche : évolution des formes, changements d'autorité ? », Communication & langages, vol. 192, no. 2, pp. 27-46.

Vuong, Q. H. (2017). Open data, open review and open dialogue in making social sciences plausible. Nature Blogs: Scientific Data. URL : <http://blogs.nature.com/scientificdata/2017/12/12/authors-corner-open-data-open-review-and-open-dialogue-in-making-social-sciences-plausible/>

Wakeling, S. et al. (2018). « No comment » ? A study of commenting on PLOS articles. Journal of Information Science. ISSN 0165-5515.

Walker, R., & Rocha da Silva, P. (2015). Emerging trends in peer review—a survey. Frontiers in neuroscience, 9, 169.

Ware, M. (2008). Peer review in scholarly journals : Perspective of the scholarly community - Results from an international study. Information Services & Use, 28(2), 109-112.

Wolfram, D., Wang, P. & Park, H. (2019). Open peer review: the current landscape and emerging models. In Proceedings of the 17th International Conference on Scientometrics & Informetrics (September 2-5, 2019, Rome, Italy).

Annexe 1 - Typologie des commentaires associés à l'évaluation de l'OPR (Wakeling et al., 2018)

PARTIE PROCÉDURALE

Commentaire du évaluateur	Copie de rapport (s) d'examineur en tant que commentaire. Comprend également les réponses des auteurs aux relecteurs et tout dialogue préalable à la publication entre le ou les auteurs et les relecteurs.
Couverture média	Commentaires rapportant ou orientant les lecteurs vers une couverture non académique / médiatique de l'article, y compris des blogs et des communiqués de presse.
Enjeu éthique	Commentaires relatifs à des conflits d'intérêts potentiels ou avérés, enquêtes sur les fautes des auteurs et les méthodes de recherche contraires à l'éthique.

Correction	Commentaires relatifs à un problème avec le texte du document et ses métadonnées, par ex. fautes de frappe, crédits manquants, légendes incorrectes, citations, etc. Cette catégorie inclut les lecteurs soulignant ces problèmes, ainsi que les éditeurs ou les auteurs notant leur correction. Mais elle n'inclut PAS les commentaires soulignant les problèmes de méthodologie de recherche, d'analyse ou d'argumentation.
Données supplémentaires	Liens ou informations sur des données supplémentaires, c'est-à-dire les données sous-jacentes ou les données justificatives de l'article.
PARTIE ACADÉMIQUE	
Critique directe	Le commentaire inclut une critique directe de l'article.
Eloge direct	Le commentaire comprend un éloge direct de l'article.
Question aux auteurs	Le commentaire pose directement une question aux auteurs.
Matériel connexe	Les commentaires incluent des liens ou des citations vers d'autres documents académiques liés à ou développant l'article en question.
Discussion	Le commentaire inclut une discussion sur le contenu du document au-delà des éloges ou critiques directs. Comprend une discussion sur les thèmes émergents de l'article, son impact potentiel, sa signification ou sa nouveauté, des suggestions d'amélioration de la recherche et la mise en évidence de problèmes méthodologiques ou analytiques perçus.
Autres	Tout commentaire ne correspondant à aucune définition de chaque catégorie

[1] Merci à Joachim Schöpfel pour sa relecture bienveillante.

[2] Peer-reviewing est ici traduit par Examen par les pairs ou Évaluation par les pairs

[3] Trois rôles : auteur-non-évaluateur, évaluateur-auteur, évaluateur-non-auteur

[4] En effet, selon les comités éditoriaux, les rejets ne sont pas toujours documentés. Ceci est cependant à pondérer selon les disciplines, avec les recommandations des sections CNU ainsi que celles de la science ouverte et ses chartes de qualité.

[5] Présenté dans la partie « Plateformes »

[6] ACP : Atmospheric Chemistry and Physics

[7] The Association of Learned & Professionnal Society Publishers.

[8] Clavey Martin. La commission européenne crée sa plateforme de publication en accès ouvert, 23 mars 2020. <https://www.soundofscience.fr/2283>

[9] <https://f1000research.com/for-authors/tips-for-finding-referees>

[10] Pour une description du fonctionnement éditorial de cette plateforme initiée par des acteurs de l'Inra : <https://peercommunityin.org/> et <https://rapidecology.com/2018/03/20/changing-the-model-of-academic-publication/>

[11] <https://fr.wikipedia.org/wiki/PubPeer> et https://fr.wikipedia.org/wiki/Catherine_Jessus

[12] <https://www.researchers.one/>

[13] Thewinnover.com

[14] La révision d'un texte est comprise dans le prix mais pas la mise-à-jour d'un article en un nouveau qui refléterait de futurs développements (400 US \$)

[15] Société de capital-investissement au capital de 6,3 milliards de dollars

[16] Voir à ce sujet, la rhétorique de la page d'accueil du service Rubriq de ResearchSquare : <https://www.researchsquare.com/company/publishers/rubriq>

[17] Analyse des résultats disponible sur : <https://dlis.hypotheses.org/5149>