

## **Guide de l'évaluation de la recherche en sciences humaines**

**Conseil sectoriel à la Recherche et à la Valorisation en Sciences humaines**

**Université de Liège**

**Septembre 2017**

Ce guide a été élaboré et rédigé par le Conseil sectoriel à la Recherche et à la Valorisation en Sciences humaines de l'Université de Liège (ULiège), sous la direction de Florence Caeymaex, Présidente et d'Ezio Tirelli, Vice-Président du Conseil :

Caeymaex Florence, F.R.S.-FNRS, Unité de Recherche (UR) Traverses  
Cloots Rudi, Vice-Recteur Recherche de l'ULiège, UR Complex and Entangled Systems from Atoms to Materials (CESAM)  
Franssen Vanessa, UR Cité  
Gautier Axel, UR Pôle de Recherche Interdisciplinaire en Sciences du Management et de l'Economie (PRISME)  
Hanocq Philippe, UR en Architecture  
Houbart Claudine, UR Art, Archéologie et Patrimoine (AAP)  
Huybrechts Benjamin, UR PRISME  
Lambert Marie, UR PRISME  
Leclercq Bruno, UR Traverses  
Leleu Yves-Henri, UR Cité  
Majerus Steve, F.R.S.-FNRS, UR Psychologie et Neurosciences Cognitives (PsyNCog)  
Martiniello Marco, F.R.S.-FNRS, UR Institut de Recherche en Sciences Sociales (IRSS)  
Michel Quentin, UR Cité  
Monseur Christian, UR Evaluation et Qualité de L'Enseignement (EQUALE)  
Moreno Paola, UR Transitions  
Oriane Jean-François, UR IRSS  
Rots Veerle, F.R.S.-FNRS, UR AAP  
Servais Véronique, UR IRSS  
Tirelli Ezio, UR PsyNCog

### **Ont contribué :**

Pirenne Vinciane, UR Mondes Anciens  
Halleux Isabelle, Administration Recherche et Développement (ARD)  
Blanchard Patricia, ARD

# TABLE DES MATIÈRES

<b>TABLE DES MATIÈRES.....</b>	<b>2</b>
<b>PRÉAMBULE.....</b>	<b>3</b>
<b>Finalités.....</b>	<b>3</b>
<b>Contexte.....</b>	<b>3</b>
<b>Principes généraux.....</b>	<b>4</b>
<b>1. DIFFÉRENTS TYPES DE TRAVAUX : SCIENTIFIQUES, D’EXPERTISE, DE VULGARISATION ....</b>	<b>5</b>
<b>1.1. Les publications scientifiques.....</b>	<b>5</b>
<b>1.2. Travaux d’expertise.....</b>	<b>5</b>
<b>1.3. Les travaux de vulgarisation.....</b>	<b>6</b>
<b>2. L’ÉVALUATION QUANTITATIVE &amp; QUALITATIVE DES PUBLICATIONS SCIENTIFIQUES.....</b>	<b>7</b>
<b>2.1. L’évaluation de la qualité extrinsèque par indicateurs quantitatifs (bibliométriques).....</b>	<b>7</b>
<b>2.2. L’évaluation de la qualité extrinsèque par indicateurs qualitatifs.....</b>	<b>9</b>
2.2.1. <i>Peer-reviewing.....</i>	<i>9</i>
2.2.2. <i>Publications ne faisant pas l’objet d’un peer-reviewing standard.....</i>	<i>9</i>
2.2.3. <i>La portée internationale des publications.....</i>	<i>10</i>
2.2.4. <i>La langue utilisée pour la publication.....</i>	<i>10</i>
2.2.5. <i>Le rang des auteurs dans les publications collectives (author credit).....</i>	<i>11</i>
<b>2.3. L’évaluation (approchée) de la qualité intrinsèque.....</b>	<b>12</b>
2.3.1. <i>Critères de la qualité intrinsèque d’une publication.....</i>	<i>12</i>
2.3.2. <i>Comment appréhender l’originalité.....</i>	<i>12</i>
2.3.3. <i>Évaluation des travaux scientifiques dans les disciplines fondées sur le recueil et le traitement de données (notamment l’inférence statistique) : qualité méthodologique.....</i>	<i>13</i>
<b>3. L’ÉVALUATION DE DEUX AUTRES TYPES D’ACTIVITÉS SCIENTIFIQUES.....</b>	<b>15</b>
<b>3.1. À propos des conférences et communications.....</b>	<b>15</b>
<b>3.2. Bases et banques de données informatiques.....</b>	<b>15</b>
<b>BIBLIOGRAPHIE INDICATIVE.....</b>	<b>16</b>

## PRÉAMBULE

### Finalités

Le Conseil Sectoriel à la Recherche et à la Valorisation en Sciences Humaines (CSRV-SH) de l'Université de Liège a élaboré un *Guide de l'évaluation de la recherche en sciences humaines* dont les finalités sont les suivantes :

- constituer un document de référence pour les pratiques d'évaluation internes à ce Conseil pluridisciplinaire en sciences humaines ;
- constituer un guide servant à orienter utilement et de manière intègre les pratiques de recherche et de publication scientifique en sciences humaines ;
- constituer un texte de référence qui positionne l'Université de Liège dans les débats nationaux et internationaux portant sur le rôle de l'évaluation dans la promotion d'une recherche scientifique *de qualité, intègre et à haute valeur d'usage scientifique et sociale*.

Ce guide reflète les orientations adoptées par le Conseil à l'éthique et à l'intégrité scientifique (CEIS) de l'ULiège, et trouve également son prolongement dans les actions de formation à la recherche menées par l'Université<sup>1</sup>.

### Contexte

Cette initiative a pour arrière-plan un certain nombre de constats. Les pratiques d'évaluation de la recherche scientifique (du *benchmarking* à l'évaluation qualitative) ne se limitent pas à faire usage d'outils de mesure ou de comparaison ; elles ont également pour effet, sinon pour but explicite, de transformer les pratiques de recherche elles-mêmes. Une certaine normalisation est l'un de ces effets structurels des pratiques d'évaluation lorsque celles-ci ont pour but d'optimiser la position d'un ou de plusieurs acteur(s) (individu, équipe de recherche, institution de recherche ou académique) dans un environnement configuré comme « compétitif ». La question de savoir si l'instauration d'environnements compétitifs est le meilleur moyen de susciter une « bonne recherche » et de favoriser l'émergence de connaissances scientifiques de haute valeur — en d'autres termes, des savoirs qui « comptent » et « font une différence » — reste ouverte et controversée. Mais, indépendamment de cette question, il est d'ores et déjà clair que beaucoup d'analyses et de débats menés à un niveau international mettent en évidence l'impact négatif qu'exercent certaines modalités d'évaluation sur les pratiques de recherche, altérant en fin de compte non seulement la nécessaire *diversité*, mais aussi la *qualité* de la production scientifique elle-même.

Dans ce contexte, il est de la responsabilité des chercheurs d'ouvrir, partout où c'est nécessaire, des espaces de discussion et de confrontation, afin de prendre une part active

---

<sup>1</sup> [http://www.ulg.ac.be/cms/c\\_6045662/fr/ethique-et-integrite-scientifique](http://www.ulg.ac.be/cms/c_6045662/fr/ethique-et-integrite-scientifique)  
<https://www.ulg.ac.be/books/formations-transversales/>

à la clarification et à l'élaboration des critères de « l'excellence » dont se réclament agences de recherche (nationales ou européennes) et institutions académiques, et que les chercheurs reprennent volontiers à leur compte, pour autant qu'il ne s'agisse pas d'un mot d'ordre et qu'ils puissent œuvrer à lui donner un sens.

Ces critères ne peuvent se réduire aux standards abstraits et figés que tendent à privilégier les dynamiques compétitives, ni à quelques indicateurs censés garantir « l'objectivité » du jugement. En matière d'évaluation, l'objectivité du jugement ne peut en effet être raisonnablement recherchée que par la confrontation et par la composition de plusieurs perspectives, elles-mêmes appuyées sur une diversité d'indicateurs et, le cas échéant, d'outils de mesure.

Pour ces raisons, un guide d'évaluation est d'abord un texte de référence et non une collection de recettes : il contient, non pas des critères « tout terrain » prêts à l'emploi, mais des principes et des points de repère à mettre en œuvre dans des situations concrètes toujours singulières.

Ces considérations ne valent pas seulement pour les humanités, dont il est admis que les modalités extrêmement diverses d'élaboration, de diffusion et de valorisation de la recherche rendent la plupart du temps impraticable et hors de propos l'usage d'outils bibliométriques standardisés : en effet, les recherches en sciences humaines qui, en psychologie par exemple, voisinent avec les sciences naturelles ou (bio)médicales rencontrent à l'heure actuelle le même type de questionnement. Ce questionnement a aussi, par conséquent, sa pertinence au-delà des sciences humaines.

Bien équipées pour élaborer un regard critique sur ce que nous font faire nos instruments de mesure, sur les rationalités qui les promeuvent et le paysage des savoirs qu'ils configurent, les sciences humaines ne défendent donc pas ici leur spécificité *contre* les autres sciences, mais la cultivent *au nom* de toutes celles qui sont mises à l'épreuve de l'évaluation.

## Principes généraux

- Assurer les conditions d'un *peer-reviewing* attentif aux qualités intrinsèques des publications ;
- Assurer l'usage d'indicateurs pertinents et variés ;
- Faire droit à la diversité des pratiques de recherche, de publication et d'évaluation dans les différentes disciplines composant le CSRV-SH (normes générales *et* déclinaisons disciplinaires) ;
- Suivre l'évolution des pratiques scientifiques et réviser les critères de l'évaluation quand c'est nécessaire ;
- À travers l'évaluation, encourager, autant que possible, une production de savoirs originaux, témoignant d'une haute exigence intellectuelle et d'un souci de partage à une échelle nationale *et* internationale.

# 1. DIFFÉRENTS TYPES DE TRAVAUX : SCIENTIFIQUES, D'EXPERTISE, DE VULGARISATION

*Les sciences humaines, tout comme les autres, peuvent déboucher sur des travaux écrits de nature et de destination diverses. Etant donné, d'une part, la diversité qui caractérise les supports de publication scientifique en SH et, d'autre part, leur évidente portée sociale (certains travaux se prêtant assez naturellement à des usages non-savants), quelques clarifications conceptuelles sont nécessaires. Distinguer les publications scientifiques « stricto sensu » (1) des travaux d'expertise (2) et de vulgarisation (3) vise à promouvoir une évaluation différenciée, sensible et adaptée aux trois modes de diffusion et de partage des savoirs en vigueur dans les SH. Si l'on attend de tout chercheur des travaux relevant incontestablement de la première catégorie, la réalisation de travaux d'expertise ou de vulgarisation peut constituer un plus, voire une priorité dans certains domaines ou dans certains contextes particuliers définis par l'Institution.*

## 1.1. Les publications scientifiques

**Une publication scientifique *stricto sensu* est une publication à destination des communautés scientifiques et, dans certains cas, d'un public spécialisé, reconnue comme contribuant à l'avancement du savoir et pour laquelle un processus de contrôle par les pairs, avant publication, est en vigueur.**

On privilégie les quatre supports de publication scientifique suivants (liste non ordonnée) :

- Articles *peer-reviewed* dans des revues,
- Ouvrages en tant qu'auteur seul,
- Ouvrages collectifs ou numéros thématiques/spéciaux de revue en tant que directeur scientifique (ou dans lequel l'auteur joue un rôle actif d'éditeur scientifique),
- Contributions et chapitres d'ouvrages collectifs (dont l' (ou les) auteur(s) joue(nt) un rôle actif d'éditeur scientifique ou prévoi(en)t des processus d'évaluation par les pairs).

Selon les disciplines, parfois fort différentes, relevant des sciences humaines, ces types de publications peuvent être diversement valorisés et hiérarchisés. Certaines disciplines valorisent particulièrement la diversification des supports de publication pour un chercheur et toutes reconnaissent la nécessité de cette diversification dans le contexte de l'évaluation d'une équipe, d'une UR ou d'une Faculté.

## 1.2. Travaux d'expertise

**Les travaux d'expertise sont des productions à destination de, et/ou répondant à une demande de publics professionnels ou spécialisés, qui reflètent la valeur d'usage des recherches qui sont à leur origine, la capacité de diffusion et de dissémination de son savoir par le chercheur et la reconnaissance sociale de celui-ci. Le *reviewing* éventuel est assuré par le commanditaire ou par d'autres experts.**

→ manuels scolaires, rapports techniques, avis experts, tests et outils de mesure, blogs spécialisés, rapports de recherche sur commande, etc.

Valoriser les travaux d'expertise revient à valoriser l'interaction entre le chercheur et son environnement non strictement scientifique, public ou privé, régional, national ou international. La nature du *reviewing* éventuel mérite d'être expliquée, si les informations sont fournies.

### 1.3. Les travaux de vulgarisation

**Les travaux de vulgarisation sont des productions écrites et orales destinées au grand public non spécialisé, qui n'impliquent pas de processus de révision par les pairs, qui ne contribuent pas à l'avancement du savoir dans le domaine scientifique concerné, mais qui soutiennent la diffusion et le partage élargis de ses résultats.**

→ ouvrages, articles dans des revues grand public, catalogues d'exposition ; *réalisation* d'émissions radio ou télévisées, de documentaires classiques ou webdocumentaires, blogs grand public, sites webs, etc. ; dans une moindre mesure, *participation* à des émissions radio, télévision, documentaires en tant qu'expert, etc.

On notera qu'il convient d'apprécier avec finesse la nature du travail engagé par le chercheur dans la vulgarisation : la *réalisation* d'un travail original, exigeant un effort de création, diffère d'une intervention ponctuelle au cours de laquelle le chercheur, en tant qu'expert, partage simplement des connaissances déjà acquises ou bien établies.

## 2. L'ÉVALUATION QUANTITATIVE & QUALITATIVE DES PUBLICATIONS SCIENTIFIQUES

*Il ne fait pas de doute que la pratique de l'évaluation en recherche se donne pour but de saisir une certaine qualité (ou valeur) de la recherche et que cette qualité n'est aucunement réductible à la seule quantité des publications d'un chercheur, d'une équipe ou d'une institution de recherche.*

*La « qualité » de la recherche est, en pratique, saisie au moyen d'indicateurs qui reflètent, soit des qualités extrinsèques d'une production scientifique (comme le nombre de citations), soit ses qualités intrinsèques, qui ne peuvent être appréciées que par l'analyse directe d'une publication (comme la solidité conceptuelle d'un raisonnement, la fiabilité d'un dispositif expérimental, l'originalité d'une hypothèse ou d'une interprétation). Nous distinguons ainsi deux grands modes d'objectivation de la qualité : l'une qui se centre sur des qualités extrinsèques, l'autre sur les qualités intrinsèques d'une ou plusieurs publications.*

*L'évaluation qui cible les qualités extrinsèques peut elle-même recourir à deux types d'indicateurs : ceux de type quantitatif (comme les indicateurs bibliométriques), et ceux de type qualitatif (comme l'autorité dont jouit une maison d'édition spécialisée, ou la qualité du reviewing dont un article a pu faire l'objet).*

*Il convient à cet égard de rappeler qu'un indicateur, qu'il soit quantitatif ou qualitatif, n'a précisément de valeur qu'indicative ou indicielle : il dénote une ou des qualités déterminées, non toutes. Son pouvoir de signification est par définition limité. « La » qualité ou « la » valeur d'une production scientifique ne serait par conséquent rien d'autre que le résultat d'une approche qui prenne en compte les limitations propres à chaque indicateur, en combinant les diverses perspectives qu'ils offrent.*

*Sur cette base, il convient de distinguer trois modalités d'évaluation :*

- l'évaluation de la qualité extrinsèque par indicateurs quantitatifs (bibliométriques),
- l'évaluation de la qualité extrinsèque par indicateurs qualitatifs,
- l'évaluation de la qualité intrinsèque.

### 2. 1. L'évaluation de la qualité extrinsèque par indicateurs quantitatifs (bibliométriques)

**Ce qu'ils indiquent sont pour l'essentiel des qualités externes ou extrinsèques comme l'influence, la notoriété ou la visibilité de certains travaux. Il s'agit par exemple :**

- du dénombrement simple d'items,
- du dénombrement des citations,
- des dénombrements relativisés (comme le H de Hirsch ou le h de Schreiber),
- du facteur d'impact (FI ou IF).

## Quelles sont les limites de ces indicateurs et des outils bibliométriques en général ?

→ Les qualités extrinsèques de notoriété ou de visibilité ne peuvent être considérées comme des traductions immédiates de la qualité intrinsèque du contenu des travaux scientifiques ; de même, les indices associés aux revues ne sont nullement garants de la qualité intrinsèque de tous les articles qui y sont publiés.

→ Beaucoup de publications SH sont mal ou non recensées par les bases de données existantes, en raison des supports de publication (papier notamment), de la politique de certains éditeurs, des systèmes hétérogènes de citation en vigueur, de la durée de vie des productions scientifiques, qui ne cadrent pas avec les durées de référence des outils bibliométriques, ou encore du privilège que ces derniers accordent aux articles dans les revues à haut index, au détriment d'autres formes de publication scientifique pouvant bénéficier d'un très grand crédit dans les disciplines SH. Nombre de productions écrites passent donc sous les radars des outils bibliométriques ou présentent des scores très faibles qui ne reflètent pas leur valeur d'usage réelle pour certaines communautés scientifiques.

→ Les outils bibliométriques sont insensibles à la très forte variabilité de l'audience relative des travaux scientifiques : ils favorisent les productions dans les domaines regroupant de très grandes communautés scientifiques, au détriment des domaines de niche, marginaux, peu financés et/ou moins à la mode, créant des effets de dominance ; l'effet est potentiellement accentué en SH où cohabitent plusieurs paradigmes, approches théoriques et démarches méthodologiques.

→ L'usage massif des indicateurs bibliométriques entremêle enjeux de performance et enjeux scientifiques. Il pousse les acteurs de la recherche à des stratégies opportunistes d'accroissement de la visibilité (multiplication, voire fragmentation, des publications, choix de thématiques, d'objets, de domaines privilégiés, diminution du temps d'élaboration de la recherche), accélère déraisonnablement le temps de la recherche, au détriment de l'originalité, de la qualité (erreurs méthodologiques, pauvreté conceptuelle, standardisation) et de la diversité (tendance à privilégier méthodologies et thématiques « facilement publiables »).

## Quelle utilisation ?

- Les indicateurs bibliométriques, mesure de la visibilité ou de la notoriété, ne peuvent constituer un mode d'évaluation, même indirect, de la qualité intrinsèque ou interne des travaux scientifiques ;
- Ces indicateurs doivent être écartés de l'évaluation dans les domaines où ils sont avérés non-pertinents et inopérants.
- Dans les domaines SH où une certaine pertinence est démontrée, ils ne peuvent constituer l'outil premier et/ou principal d'évaluation de la production scientifique ; ils peuvent dans ce contexte servir d'éléments additionnels à l'évaluation d'un projet, d'une liste de publications d'un chercheur ou d'une équipe, d'un environnement de recherche plus large, à condition d'une utilisation *proportionnée à ce qu'ils sont en mesure d'indiquer*.
- Il convient alors d'utiliser plusieurs indicateurs, en provenance de sources diverses (Scopus, Google Scholar), d'en pondérer la valeur respective et d'en contextualiser l'usage (principe de comparaison raisonnable et significative).



### **Note additionnelle à propos des classements ou référentiels de revues**

Dans la perspective d'un usage contextualisé des indicateurs, le classement de revues constitue un cas spécial.

La création concertée de référentiels intégrant et corrigeant les classements internationaux existants (ERIM, CNRS, AERES, ERIH, International Science Index) est observée en Belgique (HEC-ULg, VABB en Flandre).

Établis sur la base d'une approche multi-critères pondérée utilisant des critères à la fois quantitatifs (facteur d'impact notamment) *et* qualitatifs (tels que : qualité du *peer-reviewing* mis en place, portée internationale, langue(s) de publication, originalité et qualité de l'exécution, etc., *cf. Infra*), les classements peuvent constituer des outils pertinents d'appréciation des travaux, à condition d'intégrer des critères adaptés à la discipline ou au domaine.

On recommande dans ce cas l'usage d'une pluralité de référentiels, éventuellement avec une règle de conversion des rangs et d'arbitrage en cas de résultats divergents entre référentiels.

## **2.2. L'évaluation de la qualité extrinsèque par indicateurs qualitatifs**

**Plusieurs éléments qualitatifs sont admis comme indices de la qualité (ou valeur) d'une publication scientifique ou d'un ensemble de publications.**

### ***2.2.1. Peer-reviewing***

Il consiste en la soumission anonymisée d'articles à des évaluateurs externes au comité éditorial de la revue, ou indépendants du directeur de collection, et chargés d'en analyser la qualité, pouvant aboutir à l'acceptation de l'article sans révision, à une demande de révision (publication sous condition), ou au rejet pur et simple.

Pour l'usage de ce critère comme indicateur qualitatif, on recommande aux évaluateurs de chercher les indications les plus fiables sur la procédure de *peer-reviewing* mentionnée. Les indications fournies par le rapport généré depuis la plateforme Open Access ORBi ne sont pas complètes et il est recommandé aux auteurs eux-mêmes, comme à leurs évaluateurs, de compléter l'information le cas échéant.

### ***2.2.2. Publications ne faisant pas l'objet d'un peer-reviewing standard***

De nombreuses publications hautement valorisées en SH, parmi lesquelles les ouvrages en auteur seul, les contributions à des numéros de revue thématiques ou à des ouvrages collectifs, ne font pas l'objet d'un *peer-reviewing* standard.

D'autres indicateurs sont cependant pertinents, tels que :

- La qualité du processus de sélection, révision et correction ayant présidé à la réalisation de l'ouvrage par les auteurs. Nombre de publications sont issues de travaux collectifs antérieurs qui consistent à soumettre des hypothèses et des analyses à la critique des pairs et font, une fois écrites, l'objet d'un second

*reviewing* non anonyme visant à la révision du manuscrit ; d'autres, sollicités, font directement l'objet d'un *reviewing* par l'éditeur scientifique et son équipe, pouvant aboutir lui aussi à la révision du manuscrit. Aussi doit-on faire une nette différence entre ce type de publication, d'élaboration longue et exigeante, et une simple collection d'abstracts ou de communications (issus d'un colloque par exemple).

- La maison d'édition (prestige, reconnaissance scientifique de son catalogue comme référence pour tel ou tel domaine, qualité connue de ses procédures de sélection et de révision).
- Les citations dont l'ouvrage ou la contribution fait l'objet dans des travaux de référence.
- Les comptes rendus dont l'ouvrage ou la contribution fait l'objet.

On ne peut qu'encourager les déposants à donner ces informations, notamment en offrant des possibilités de « commentaires additionnels » dans les choix ou listes de publications figurant dans les dossiers.

### **2.2.3. La portée internationale des publications**

On s'accorde à reconnaître l'importance, pour un chercheur ou une équipe, d'assurer la diffusion et le partage de ses recherches et de ses résultats à un niveau, non seulement national, mais international.

Comme indicateurs de la portée internationale, on peut prendre en compte :

- pour un article ou un ouvrage : le caractère international de l'équipe éditoriale de la revue ou de la collection ;
- pour un article : le caractère international des *reviewers* quand on peut l'objectiver ;
- pour un dossier scientifique : la stratégie de diffusion internationale du chercheur ou de l'équipe (langues utilisées, diversité des supports de publication) ;
- pour une contribution à un ouvrage ou à un numéro thématique de revue : l'affiliation des autres auteurs.

Il faut toutefois prendre en considération le fait que :

- certaines revues nationales ont un impact international (ex. : *British Journal of Sociology*, *Revue Française de sociologie*) ;
- certaines revues dites internationales ont une diffusion et une reconnaissance limitée ;
- les SH ont parfois affaire à des objets locaux, géographiquement et culturellement situés, suscitant l'intérêt des communautés concernées et de communautés scientifiques spécialisées et plus localisées (ex : linguistique néerlandaise, musicologie des minorités culturelles, droit fiscal belge, histoire de Belgique).

### **2.2.4. La langue utilisée pour la publication**

Dans la perspective d'une diffusion et d'un partage élargis des connaissances, la capacité du chercheur à publier certains de ses travaux dans des langues différentes de la sienne constitue un plus certain, voire une nécessité dans certaines disciplines (telles la finance ou certains sous-domaines de la psychologie).

L'anglais offre aujourd'hui à l'auteur en SH la possibilité la plus large de diffuser son travail (notamment en direction des pays de langues minoritaires, où la pratique de l'anglais comme seconde langue est une condition *sine qua non* de l'intégration à une communauté scientifique) et les publications en anglais méritent d'être encouragées et valorisées, notamment comme moyen de faire vivre et de soutenir les traditions, styles de pensée, paradigmes et types d'approches minorisées ou négligés dans le monde scientifique anglo-saxon.

Concernant l'usage des autres langues, il convient cependant de noter que :

- le français, toujours en expansion pour des raisons démographiques, reste soutenu par d'importants investissements institutionnels et commerciaux qui continuent de lui assurer le rôle d'une langue scientifique et intellectuelle à l'échelle internationale ;
- certains domaines spécifiques, pour des raisons historiques, culturelles et géographiques, peuvent privilégier d'autres langues au titre de *lingua franca* ou langue de référence (par exemple l'italien pour ce qui concerne les études sur l'Italie renaissante, l'allemand pour l'archéologie ou l'histoire de l'art, le français pour la littérature et la philosophie contemporaines françaises, le français dans les études de droit continental par opposition à la *common law* anglo-saxonne) ;
- s'il s'agit d'accroître la diffusion et la visibilité à un niveau international, d'autres langues offrent des possibilités très importantes : l'espagnol par exemple.

En conclusion, les publications en anglais ne constituent pas le seul indicateur, ni nécessairement le plus pertinent, pour apprécier le profil international d'un chercheur. L'évaluation doit prendre en considération la diversité des langues dans lesquelles le chercheur se montre capable de publier, en relation avec les spécificités du domaine et le public scientifique visé.

### **2.2.5. Le rang des auteurs dans les publications collectives (author credit)**

Dans les SH, la publication en auteur seul ou à nombre restreint d'auteurs (2 à 3) reste la pratique dominante. Toutefois, certains domaines ont adopté la pratique des publications à plusieurs auteurs ; si celle-ci doit être encouragée comme favorisant collaborations et travail d'équipe, elle pose le problème de l'importance relative du travail consenti par chacun des auteurs.

Afin de décourager les usages stratégiques (maximisation du nombre de publications), on doit promouvoir, soit l'usage standardisé de rangs (premier, dernier auteur), soit l'explication par l'auteur du rôle précis joué dans la recherche et la publication. L'évaluateur doit y accorder une attention particulière lorsque l'information est accessible.

## 2.3. L'évaluation (approchée) de la qualité intrinsèque

Les limites normales de l'approche par indicateurs des qualités extrinsèques (quantitatifs comme qualitatifs) et le fait que l'information disponible est parfois partielle sont des faits qui plaident en faveur d'une évaluation tournée vers la qualité intrinsèque, basée sur la lecture des publications scientifiques. Cette méthode présente ses propres critères, relatifs au contenu de la publication. Nécessitant l'intervention d'évaluateurs proches du domaine, elle n'exclut pas, toutefois, l'appréciation de *reviewers* plus éloignés, voire généralistes.

Cette méthode, reposant sur le principe du *peer-reviewing*, doit être appliquée, autant que possible, pour les évaluations individuelles des chercheurs de niveau avancés et/ou porteurs de projets à forts enjeux financiers (postdoctorants, promotions académiques, projets portés par des chercheurs confirmés).

Sa mise en œuvre suppose :

- de généraliser le principe d'une sélection par le chercheur (ou une équipe) d'un choix de publications représentatives, avec possibilité de commentaire additionnel sur les raisons de ce choix et le contexte d'élaboration de la publication ;
- de garantir aux évaluateurs un accès immédiat à ces publications ;
- de prévoir l'obligation, pour les évaluateurs, de rendre compte de leur lecture, sur la base de critères définis (voir ci-dessous 2.3.1)
- de garantir aux évaluateurs un temps de travail raisonnable pour l'évaluation, en opérant des choix stratégiques, privilégiant la qualité et l'utilité, quant au volume global du travail d'évaluation (principe du *less is more*).

### 2.3.1. Critères de la qualité intrinsèque d'une publication

- **Théorie** : bonne information, solidité des bases conceptuelles et théoriques, pertinence ou originalité de la question, du mode de problématisation ou de l'hypothèse de recherche par rapport à l'évolution de la discipline, qualité de l'interprétation des résultats obtenus, éventuelle innovation paradigmatique par rapport au champ.
- **Méthodologie** : pertinence des outils, clarté de la méthode, rigueur conceptuelle et/ou expérimentale, originalité et modalité de recueil des données dans le cas d'études empiriques, originalité de la méthodologie.
- **Aspects formels** : clarté de l'écriture, de la présentation et de l'explicitation pour les généralistes.

### 2.3.2. Comment appréhender l'originalité

L'originalité, très valorisée sur le plan théorique et méthodologique, est difficile à appréhender. Plusieurs indicateurs s'offrent aux évaluateurs, qui doivent être autant que possible proches du domaine :

- Le positionnement explicite de la publication par rapport à l'état de l'art ;
- La variété des travaux d'un chercheur ou d'une équipe ; multiplicité des objets, des méthodologies, des postures épistémologiques ou des formats de publication peuvent être envisagées comme signes de développement, de diversification, de recherche de connaissances nouvelles.

L'un des meilleurs moyens reste la soumission à des experts externes, explicitement mandatés pour effectuer exclusivement l'évaluation des qualités intrinsèques, l'évaluation des qualités extrinsèques (quantitatives et qualitatives) étant effectuée par l'institution.

### ***2.3.3. Évaluation des travaux scientifiques dans les disciplines fondées sur le recueil et le traitement de données (notamment l'inférence statistique) : qualité méthodologique***

Certains domaines, dont les résultats reposent sur le recueil et le traitement de données, connaissent actuellement une crise de « reproductibilité » des résultats, touchant le socle épistémologique de ces disciplines. Les mouvements associés à la méta-recherche et à l'*Open Science*, plaident pour un changement dans les pratiques ; ils proposent pour ce faire de transformer :

- les critères et modalités d'évaluation des travaux publiés, en prenant en compte la qualité méthodologique (a) ;
- la production d'articles elle-même, en y intégrant une évaluation par les pairs de premier niveau (b).

#### *(a) Critères et modalités d'évaluation méthodologique*

1. Usage ou non usage d'outils statistiques permettant de présager une reproductibilité statistique, tels que le calcul des effectifs ou de la puissance statistique ;
2. Pratique du double-aveugle, utilisation (ou pas) de blocs randomisés.

Aux auteurs comme aux évaluateurs, on conseille de se reporter aux *guidelines* élaborés par domaines. Voir le site EQUATOR d'Oxford : PRISMA (méta-analyses) ; ARRIVE (recherche sur modèles animaux) ; CONSORT (expériences avec participants humains)

#### *(b) Évaluation par les pairs, intégrée au processus de production des publications*

1. Pré-enregistrement du protocole d'investigation sur une plateforme avant la réalisation de l'étude ; celui-là aura le statut d'article publié en *Open access* et sera citable. Les auteurs y feront référence (avec éventuelle discussion) dans l'article qui en découlera.
2. Partage (*data sharing, open data, open methods, open tools*) des données brutes, analyses et de certaines méthodes sur un site spécialisé (ou un journal scientifique) ; documents portant le nom des auteurs et citables ; partage à signaler dans l'article publié en aval. La démarche vise notamment à encourager la transparence et la réalisation de méta-analyses qui fournissent « le plus haut degré de preuve » au sujet d'un effet, défavorisant le biais de publication, et favorisant la reproduction/reproductibilité des résultats par d'autres chercheurs.
3. Pré-enregistrement ou pré-exposition du manuscrit complet dans un journal électronique (ex. : *archBioxiv*) citable, avant sa soumission à un « vrai » journal. Une telle démarche permet de recevoir (potentiellement) des critiques et des suggestions de *peer-reviewers* non-sélectionnés et libres.

4. Publication d'études visant la réplication et la reproductibilité d'autres études déjà visibles dans le domaine. Il s'agit de l'une des démarches de la méta-recherche elle-même. L'agence gouvernementale de recherche des Pays-Bas a réservé (2017) une part importante de son budget à des projets de reproductibilité (permettant par exemple de confirmer si une étude apparemment novatrice l'est vraiment).
5. Implication explicite et sans « complaisance » d'un méthodologue dans l'étude, en tant que véritable co-auteur.
6. Participation à des consortia de recherche. Il est démontré que la qualité des recherches publiées par ce type de groupes de chercheurs (parfois intercontinentaux) est d'une qualité très supérieure à celle de la moyenne (petits groupes). Ce critère ne concerne pas toutes les disciplines.
7. Participation du chercheur ou du centre de recherche à un *Open science framework*.

Ces recommandations ne s'adressent pas seulement aux chercheurs en tant qu'auteurs, mais aux évaluateurs.

Les différents types de publications Open peuvent évidemment constituer un matériau de choix pour l'évaluation de la qualité intrinsèque des travaux, entre autres dans leurs aspects méthodologiques. Mais on peut aussi envisager d'intégrer le simple usage des *Open tools*, par les auteurs, au nombre des indicateurs qualitatifs positifs de qualité extrinsèque, en bref : d'en récompenser l'usage. Rendre publique l'utilisation de l'*Open* (-science, -data, -tools) dans l'évaluation encouragera les chercheurs eux-mêmes à y recourir davantage.

La promotion de l'*Open Science* comme moyen de régulation des pratiques de publication et d'amélioration de l'évaluation s'inscrit dans un large mouvement, encouragé par la Commission européenne et de nombreux autres acteurs de la recherche<sup>2</sup>.

---

<sup>2</sup> <https://www.openaire.eu/>  
<http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-policy-platform>

## 3. L'ÉVALUATION DE DEUX AUTRES TYPES D'ACTIVITÉS SCIENTIFIQUES

### 3.1. À propos des conférences et communications

Sans être des publications au sens strict, les conférences et les communications sont valorisées comme autant de démarches de communication (diffusion et partage) des résultats de la recherche, de confrontation critique aux pairs de la discipline, mais aussi de reconnaissance par ces derniers.

Trois critères majeurs sont pris en compte dans leur appréciation :

1. modalité de sélection : invitation, invitation comme *keynote*, sélection rigoureuse suivant un *call for papers* ;
2. portée internationale de la manifestation : (*cf. supra* pour la portée internationale et la langue) à distinguer de l'ampleur de la manifestation ;
3. distance relative avec l'équipe locale d'une partie au moins des communications scientifiques.

### 3.2. Bases et banques de données informatiques

Les bases ou banques de données en ligne à destination des chercheurs (*Digital Humanities, Humanités numériques*) ne sont pas toujours des publications scientifiques au sens défini plus haut, notamment lorsqu'elles ne font pas l'objet d'un processus de sélection et de correction par les pairs.

Leur valeur tient à :

- leur pertinence pour une ou des communautés scientifiques données et devrait s'appréhender par des informations concernant leur destination et leur usage ;
- la nature du travail qui préside à leur élaboration (la simple mise à disposition de corpus par numérisation doit être distinguée de l'élaboration d'outils spécifiques, ou de l'organisation complexe de l'information)

Des indications pertinentes à faire apparaître dans un dossier, au travers de « commentaires additionnels » et à prendre en compte par les évaluateurs seront :

- La contribution effective du chercheur au travail réalisé. Ces bases de données sont le plus souvent collaboratives et elles requièrent des contributions et des compétences distinctes : conception, développement technique, alimentation en données, développement de métadonnées et d'outils d'exploitation, etc.
- Les éventuels indices d'usage par des chercheurs et/ou publics spécialisés (citations dans des publications, résultats théoriques ou appliqués, etc.).



## BIBLIOGRAPHIE INDICATIVE

- Alberts B., (2013) « Impact factor distortions ». *Science*, 340, p.787.
- Allen L. et al., (2014) « Credit where credit is due ». *Nature*, 508, p. 312-314.
- Benedictus R., Miedema F., (2016) « Fewer numbers, better science ». *Nature*, 538, p. 453-455.
- Biagioli M., (2016) « Watch out for cheats in citation game ». *Nature*, 535, p.201.
- Bersini H., Clerbaux B., Hudon M., Leclercq B., Timmermans B., Zaccai E. (eds.), (2016) *L'évaluation de la recherche en question*. Bruxelles, Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique.
- Brembs B., et al., (2013) « Deep impact: unintended consequences of journal rank ». *Frontiers in Human Neuroscience*, 7, article 291.
- Bruno I., (2008) « La recherche scientifique au crible du benchmarking. Petite histoire d'une technologie de gouvernement ». *Revue d'histoire moderne et contemporaine*, 55, 4bis, p. 28-45.
- Bruno I., (2009) « Le temps des "chercheurs-entrepreneurs": sens et pouvoir du benchmarking dans l'"espace européen de la connaissance" ». *Quaderni*, 69, p. 93-104.
- Bureau M.-C., (2010) « Du travail à l'action publique : quand les dispositifs d'évaluation prennent le pouvoir ». *Cahiers internationaux de sociologie*, 128-129, p. 161-175.
- Büttgen P., Cassin, B., (2009) « J'en ai 22 sur 30 au vert. 6 thèses sur l'évaluation ». *Cités*, 37, 1, p. 27-41.
- Callaway E., (2016) « Publishing elite turns against impact factor ». *Nature*, 535, p. 2010-211.
- Chambers C. D., (2017) *The seven deadly sins of psychology. A manifesto for reforming the culture of scientific practice*. Princeton, NJ, Princeton University Press.
- Colletis G., Lamarche T., (2011) *Pour une nouvelle démarche d'évaluation des laboratoires de recherche, des enseignants-chercheurs et des chercheurs en économie*. Rapport de la Commission d'évaluation de l'Association française d'économie politique ([http://assoekonomiepolitique.org/pour\\_une\\_nouvelle\\_demarche\\_devaluation\\_des\\_laboratoires\\_de\\_recherche\\_des\\_enseignants-chercheurs\\_et\\_des\\_rechercheurs\\_en\\_economie/](http://assoekonomiepolitique.org/pour_une_nouvelle_demarche_devaluation_des_laboratoires_de_recherche_des_enseignants-chercheurs_et_des_rechercheurs_en_economie/)).
- Cronin B., Sugimoto C. R., (2014) *Beyond bibliometrics. Harnessing multidimensional indicators of scholarly impact*. Cambridge, MA, MIT Press.
- Debackere K., Glänzel W., (2004) « Using a bibliometric approach to support research policy making: the case of the Flemish BOF-key ». *Scientometrics*, 59, 2, p. 253-276.



De Bellis N., (2009) *Bibliometrics and citation analysis. From the science citation index to cybermetrics*. Lanham, MD, Scarecrow Press.

Desrosières A., (2008) *Gouverner par les nombres*. Paris, Presses des Mines ParisTech.

Destro Bisol G. et al. (2014) « Perspectives on Open Science and scientific data sharing: An interdisciplinary workshop ». *Journal of Anthropological Sciences*, 92, p. 179-200.

American Society for Cell Biology, (2013) *DORA: San Francisco Declaration on Research Assessment: Putting science into the assessment of research*. <http://www.ascb.org/dora/>.

Franco A. et al., (2014) « Publication bias in the social sciences: unlocking the file drawer ». *Science*, 345, p. 1502-1505.

Gingras Y., (2014) *Les dérives de l'évaluation de la recherche. Du bon usage de la bibliométrie*. Paris, Editions Raisons d'agir.

Gingras Y., (2016) *Bibliometrics and Research Evaluation. Uses and Abuses*, Cambridge, MIT Press.

Greenberg S. A., (2009) « How citation distortions create unfounded authority: analysis of a citation network ». *BMJ*, 339, p. b2680.

Harris R., (2017) *Rigor mortis. How sloppy science creates worthless cures, crushes hope and wastes billions*. New York, Basic Books.

Hicks D. et al., (2015) « The Leiden manifesto for research metrics ». *Nature*, 520, p. 429-431.

Higginson A. D., Munafo M. R., (2016) « Current incentives for scientists lead to underpowered studies with erroneous conclusions ». *PLOS Biology*, 14, p. e2000995.

Ioannidis J. P., (2005) « Why most published research findings are false ». *PLOS Medicine*, 2, p. e124.

Ioannidis J. P., (2014) « How to make more published research true ». *PLOS Medicine*, 11, p. e1001747.

Ioannidis J. P. et al., (2015) « Meta-research: evaluation and improvement of research methods and practices ». *PLOS Biology*, 13, p. e1002264.

Ioannidis J. P. et al., (2016) « Citation metrics: A primer on how (not) to normalize ». *PLOS Biology*, 14, p. e1002542.

Kreiner G., (2016) « The slavery of the h-index: Measuring the unmeasurable ». *Frontiers in Human Neuroscience*, 10, art. 556.

Lan J., (2010) « Let's make science metrics more scientific ». *Nature*, 464, p. 488-489.

McKiernan E. C. et al., (2016) « How open science helps researchers succeed ». *eLife*, 5, p. e16800.

Munafò M.R. et al., (2014) « Scientific rigor and the art of motorcycle maintenance ». *Nature Biotechnology*, 32, p. 871-873.

Munafò M.R. et al., (2017) « A manifesto for reproducible science ». *Nature Human Behaviour*, 1, art. 0021, p. 1-9.

Sahel J. A., (2011) « Quality versus quantity: Assessing individual research performance ». *Science Translational Medicine*, 3, p. 84cm 13.

Sarewitz D., (2016) « The pressure to publish pushes down quality ». *Nature*, 533, p. 147-148.

Schekman R., (2013) « How journals like Nature, Cell and Science are damaging science ». *The Guardian*, 9 December 2013.

Stephan P., Veugelers R., Wang J., (2017) « Blinkered by bibliometrics ». *Nature*, 544, p. 411-412.

Vinokur A., (2006) « La qualité de la mesure de la qualité dans l'enseignement supérieur : essai d'analyse économique ». *Éducation et sociétés*, 18, p. 109-124.

Waltman L., van Eck N. J., (2012) « The inconsistency of the h-index ». *Journal of the American Society for Information Science and Technology*, 63, p.406-415.

Weingart P., (2005) « Impact of bibliometrics upon the science system: Inadvertent consequences? ». *Scientometrics*, 62, 1, p. 117-131.