

## Trop d'évaluations tuent l'évaluation ?



Rien de tel que de passer quelques jours dans l'atmosphère d'un colloque au sein d'un institut où les chercheurs s'activent, discutent comme dans une ruche bruisant de sciences et d'innovations, pour retrouver l'enthousiasme de nos jeunes années. Cela n'empêche pas la réflexion sur l'évolution de la recherche au cours de ces dernières décennies, et surtout de son évaluation.

C'est en discutant avec quelques enseignants-chercheurs ayant bénéficié d'«advanced grant ERC» (bourses de l'European Research Council) et qui manifestement s'éclataient dans une unité disposant de moyens de pointe que me sont revenues les paroles de Carlos Moedas lors de l'anniversaire des dix ans du Conseil européen de la recherche à Strasbourg en 2017 : « *La beauté de ce programme est d'aller chercher les meilleurs scientifiques pour leur demander ce qu'ils veulent faire plutôt que ce qu'ils doivent faire.* » Inutile de dire que dans cet esprit, les candidatures européennes n'ont pas manqué et que la sélection, même sur programmes libres, reste sévère.

Ceci contraste cependant avec nombre d'appels d'offres européens, nationaux ou régionaux, où la liberté d'entreprendre est fortement bridée par des thématiques ciblées qui sont souvent les mêmes et d'une uniformité

désespérante. Les armées d'évaluateurs ont à leur disposition presque toujours les mêmes grilles: qualité scientifique (efficacité et innovation), résultats déjà obtenus (bibliométrie), à noter suivant la même hiérarchie (excellent, très bon, bon, marginal, mauvais).

Il en est de même pour l'évaluation des établissements, des unités et laboratoires. Car depuis une vingtaine d'années, le nombre d'évaluations s'est multiplié, et qu'à côté du Comité national de la recherche scientifique (CoNRS) et du Conseil national des universités (CNU), qui concernaient surtout l'évaluation et la carrière des chercheurs et enseignants, s'est développée une culture d'évaluation des structures. Cette (r)évolution copernicienne est issue d'une logique de mission régalienne que sont la recherche et l'enseignement supérieur pour l'État. Des indicateurs de suivi des budgets consacrés à cette mission doivent être mis en place pour éclairer le gouvernement et le parlement sur les performances des politiques suivies et sur « *ce que font ces foutus chercheurs budgétivores.* »

D'où la création de l'Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES), qui a mené des évaluations prescriptives et qui est très vite tombée dans le « benchmarking » généralisé de l'approche comptable

de la recherche et ses dérives avec trop de références bibliométriques, la dictature du nombre de citations (SCI Thomson, Web of Science...) et du facteur h. Sont alors apparus comme deux icônes le mythe de l'excellence et celui de la qualité, bien que je n'ai pas encore trouvé une définition claire de ces icônes. Il ne s'agissait plus alors d'être bon mais d'être le meilleur pour gagner quelques places dans le classement de Shanghai, toujours décrié mais souvent consulté.

Pour la politique imaginative des ministères durant les années 2000, on parle d'une « fièvre de l'évaluation » qui, d'après certains observateurs, est plus contagieuse que certaines gripes. Car avec le nombre d'évaluations, c'est le nombre d'évaluateurs qui est frappé d'inflation ! Aux 1 000 membres du CoNRS et 3 500 membres du CNU habituels se sont ajoutés 2 800 experts du HCERES (Haut conseil de l'évaluation de la recherche et de l'enseignement supérieur, qui a succédé à l'AERES en 2013), managés par 119 délégués scientifiques et 110 administratifs, puis les 1 650 experts de l'ANR, et les 876 membres des CES (Comités d'évaluation scientifique) administrés par 283 collaborateurs. Il est vrai qu'avec le nombre de dossiers traités, accessibles dans les rapports 2016 de ces organismes, c'est un travail titanesque qui impose une importante fraction à plein temps pour faire face à l'évaluation des établissements vague après vague, celle des programmes et projets imaginés par le ministère et la SNRI (Stratégie nationale de recherche et d'innovation) – IDEX, LABEX, Programmes d'investissements d'avenir (PIA 1, 2 et 3) –, sans compter le dernier avatar : les écoles universitaires de recherche (« Graduate Schools »). Tant et si bien que des voix s'élèvent : « *Trop d'évaluations tuent l'évaluation.* »

Le mal n'est pas seulement français mais européen ; pour preuve, un très bon texte de trois académies européennes – l'Académie des sciences, la Royal Society et la Leopoldina – paru en octobre 2017 [1] qui sonne comme un manifeste et donne à son tour un avis pertinent sur la pratique de l'évaluation des programmes de recherche et des chercheurs. Les trois académies rappellent en particulier le mésusage du facteur d'impact (« impact factor », IF) et le recours dans l'évaluation du nombre de publications, et qui plus est dans les journaux à fort indice, conduisant à la pratique critiquable de donner trop de poids aux indicateurs bibliométriques. Elles notent que des travaux importants et originaux ont aussi été trouvés dans des journaux à faible facteur d'impact. La bibliométrie conduit les chercheurs à publier dans les sujets à la mode en négligeant les « chemins escarpés », attitude qui leur apporte force de citations, score qu'ils peuvent encore améliorer grâce aux « citations clubs » ou groupes de chercheurs qui se citent mutuellement. Elles dénoncent pour l'évaluation individuelle

la tyrannie du facteur h et la surcotation des impacts par les mesures « altimétriques ». En se penchant sur l'évaluation par les pairs, elles pensent qu'il est crucial que les évaluateurs qui doivent apprécier la qualité et l'originalité d'un dossier soient aussi eux-mêmes de haut niveau, sinon rangés parmi les leaders dans leur discipline. Dans cette optique, il est clair que le nombre d'excellents évaluateurs est limité, et donc que le nombre de processus d'évaluation doit être réduit sous peine de manquer d'évaluateurs de première classe !

L'ensemble des évaluateurs experts est une ressource qui par devoir académique est souvent surexploitée dans un nombre croissant de comités chronophages et peu productifs. Les académies donnent quelques conseils ou pistes d'amélioration : les soumissions doivent se faire en un nombre de pages limité, l'évaluation a besoin de se concentrer sur l'essentiel, les dossiers excessivement longs étant contre-productifs ; la rotation des évaluateurs doit s'exercer pour renouveler le panel des experts, qui doit aussi refléter la diversité des champs scientifiques. Enfin, elles insistent sur les principes d'éthique et soulignent que la compétence des « reviewers » et des évaluateurs n'est pas innée ; elle s'apprend par des méthodes et des apprentissages qui évitent les « biais inconscients » par une approche impartiale et expérimentée. Dans leur conclusion, les trois académies répètent que les comités d'évaluation demandent des experts connus et reconnus aux plus hauts standards éthiques, concentrant leurs avis sur les mérites intellectuels et les portées des résultats scientifiques. Les données bibliométriques ne peuvent remplacer un jugement bien fondé, et donner trop d'importance à la mesure des publications peut sérieusement faire passer à côté de la créativité et de l'originalité scientifique.

Nous avons déjà discuté ici du fameux « publish or perish », qui est une pire dérive pour le monde académique et celui de la recherche. Les scientifiques soumis à la pression sont plus incités à produire plutôt qu'à découvrir, d'où la « junk science » [2]. Il me revient que Peter Higgs, le physicien père du fameux « boson » qui porte son nom, rappelait lors de la remise de son prix Nobel en 2013 qu'il ne pourrait plus faire le même type de recherche aujourd'hui, car il ne trouvait pas assez vite et ne publiait pas assez. L'Université d'Edimbourg était à deux doigts de se passer de ses services lorsqu'il fut pour la première fois nommé dans les années 1980.

**Jean-Claude Bernier**

Mai 2018

[1] Statement by three national academies (Académie des Sciences, Leopoldina and Royal Society) on good practice in the evaluation of researchers and research programmes, 27 oct. 2017, [www.academie-sciences.fr/pdf/rapport/avis111217.pdf](http://www.academie-sciences.fr/pdf/rapport/avis111217.pdf)

[2] Bernier J.-C., Y a-t-il une « junk science » ?, *L'Act. Chim.*, 2016, 403, p. 5.

