

L'expertise scientifique et le choix des experts en question

Jacques Reisse

Professeur Emérite de l'Université Libre de Bruxelles

Membre de l'Académie royale de Belgique

Résumé

L'expertise scientifique joue un rôle majeur dans nos sociétés puisque les décisions politiques ou économiques, pour ne citer qu'elles, sont prises sur base d'avis d'experts. Il est donc intéressant de s'interroger sur le fonctionnement même de l'expertise, sur les règles auxquelles elle doit satisfaire pour être crédible et surtout sur les critères de choix des experts. La tendance croissante de sélectionner les experts, comme d'ailleurs aussi les chercheurs, sur base du nombre de leurs publications et du taux de citation de ces publications conduit à des absurdités qu'il convient de dénoncer avec force. En effet, lorsque les critères de choix des experts sont critiquables, l'expertise scientifique perd de sa crédibilité. Une telle perte de crédibilité pourrait, à terme, avoir des conséquences graves sur le fonctionnement de nos sociétés démocratiques.

Introduction

Les mots « expertise » et « experts » résonnent à nos oreilles plusieurs fois par jour. Qu'il s'agisse d'une crise politique ou économique, de l'issue possible d'un conflit, de l'évolution du climat ou plus prosaïquement du temps prévu pour le prochain week-end, nous attendons l'avis des experts. Cet avis nous est parfois transmis directement mais le plus souvent il nous arrive de manière indirecte par la voie des médias ou au travers de déclarations de décideurs qui avaient sollicités l'avis d'experts pour fonder leur décision.

En effet, les décideurs, qu'ils soient membre d'un gouvernement, banquier, chef d'entreprise ou guide de montagne prêt à mener des clients au sommet du Cervin, prennent très généralement leurs décisions sur base de tels avis. L'expertise est une chose sérieuse qui conditionne notre vie individuelle mais aussi celle de la cité et il est normal donc de s'intéresser à son fonctionnement, à ses règles quand elles existent, à ses limites. C'est ce que nous allons faire dans cet article.

Les différents types d'expertise et la relation expert-décideur

On demande aux experts de dire si, selon eux, tel tableau est une œuvre originale ou un faux, si l'assassin est sain d'esprit ou s'il doit être placé dans un établissement psychiatrique plutôt que dans une prison ou encore, si un pont nouvellement construit l'a été

conformément au cahier des charges. Dans tous ces cas, il s'agit d'une expertise d'évaluation.

On demande aussi aux experts s'il existe un risque que tel volcan entre très prochainement dans une phase éruptive cataclysmique ou encore, si le décrochement observé dans la faille de San Andrea peut conduire à un tremblement de terre majeur dans l'année qui vient. On demande à l'expert quelles pourraient être les conséquences sur le climat de l'Europe occidentale à l'horizon 2030, du maintien, au niveau actuel, de la production annuelle de CO₂. Dans ces trois derniers cas, il s'agit d'une expertise qui conduit à des prédictions concernant le futur et c'est précisément ce type d'expertise qui va nous intéresser dans les pages qui suivent.

La complexité des phénomènes que nous venons d'évoquer (volcanisme, tectonique, climat) est telle qu'il serait absurde de demander à l'expert et plus généralement à un groupe d'experts de n'exprimer leur avis que s'ils sont absolument certain de la justesse de leurs appréciations. Une telle demande mettrait les experts sérieux, ceux qui sont de vrais scientifiques, à se retirer du groupe. En revanche, il est licite de demander aux experts d'estimer le degré de validité de leurs avis en introduisant, dans leur rapport, un élément probabiliste. A titre d'exemple, nous savons que toutes les prévisions météorologiques professionnelles sont assorties d'un tel facteur.

Le décideur, quant à lui, doit donc transformer cet avis « nuancé » en une décision qui, elle, ne l'est pas. Si le risque d'une éruption cataclysmique ne peut être écarté et même si les experts estiment la probabilité de l'évènement à 10% (ou si l'on préfère à une « chance » sur dix), il est probable que le pouvoir politique décidera de l'évacuation de la population qui vit dans la zone à risques. Si la probabilité que survienne un tremblement de terre majeur dans une zone surpeuplée est estimée à 10%, le pouvoir politique décidera très probablement d'informer la population en donnant des consignes précises quant aux comportements à adopter si le phénomène a lieu. Si le maintien de la production actuelle de CO₂ peut provoquer, avec une probabilité de 50% une élévation significative du niveau des mers et une augmentation importantes du nombre de tempêtes à l'horizon 2030, le pouvoir politique n'aura pas d'autre choix que de prendre la décision de renforcer les digues et d'évacuer les personnes habitant des zones inondables.

Le décideur fonctionne nécessairement dans une logique binaire de type 0/1 : il prend ou on ne prend pas telle ou telle décision. En revanche, la logique de l'expert sollicité pour donner son avis sur des questions d'une extrême complexité est une logique à valeurs multiples ce qui signifie qu'entre 0 et 1, existent toutes les valeurs intermédiaires.

Définition de l'expertise scientifique

Nous adopterons ici la définition proposée en 1999 par Philippe Roqueplo qui a consacré plusieurs articles et ouvrages à l'expertise scientifique.

Pour Philippe Roqueplo, l'expertise scientifique est :

« l'expression d'une connaissance formulée en réponse à une demande de ceux qui ont une décision à prendre en sachant que cette réponse est destinée à être intégrée dans un processus de décision »

Pour compléter cette définition et toujours selon Philippe Roqueplo :

« Entre savoir et décision : l'expertise scientifique »

Il s'agit là du titre de l'ouvrage que cet auteur a publié en 1996 (Edition de l'INRA) et dans lequel, il analyse avec beaucoup de lucidité l'ambiguïté que peut revêtir le terme d'expert. Ainsi, l'expert envoyé par son gouvernement pour le représenter est certes choisi pour ses connaissances du sujet mais il l'est aussi pour son aptitude à défendre les intérêts de son pays. En revanche, l'expert choisi pour éclairer le décideur sur une décision à prendre doit sa sélection, en principe du moins, à sa seule compétence scientifique. Dans le cadre de cet article, nous nous intéresserons exclusivement à ce type d'expert.

Retenons de la définition de l'expertise scientifique qu'il existe un demandeur d'expertise et des experts et que leurs rôles sont clairement différents. Les experts formulent leurs conclusions en sachant l'usage qui va en être fait mais la décision quant aux mesures à prendre est de l'entière responsabilité du demandeur d'expertise. Il s'agit là d'une situation idéale qui, malheureusement, est rarement réalisée.

Ainsi que nous l'avons dit précédemment l'expertise scientifique est une activité d'une extrême importance et il est normal donc qu'elle soit soumise à des normes. Prenons comme exemple la charte de l'expertise scientifique collective telle qu'elle a été adoptée par l'Institut National de Recherche Agronomique (INRA) français. Cet institut de recherche agronomique public finalisée est le plus important d'Europe et le second au niveau mondial. Les activités de recherche et d'expertise de l'INRA couvrent l'agriculture, la sécurité alimentaire, l'environnement et d'autres domaines connexes ; les normes définies par cette institution ont donc valeur d'exemple pour notre propos.

Selon les normes de l'INRA, l'expertise scientifique est fondée sur quatre piliers qui sont :

- la compétence des experts
- la pluralité et donc la diversité des approches des experts
- l'impartialité des experts
- la stricte séparation des rôles entre experts et demandeur d'avis

L'INRA précise la procédure à suivre pour s'assurer de la compétence des experts:

« la compétence des experts est garantie par le mode de repérage et de sélection des experts qui s'opère sur base de leurs publications validées par leurs pairs »

On serait tenté de considérer qu'il s'agit là d'un critère de repérage et de sélection qui ne devrait mériter aucun commentaire additionnel puisque c'est ainsi que les organismes de recherche, qu'ils soient publics ou privés, procèdent depuis longtemps pour sélectionner les chercheurs. En réalité, et ainsi que nous le verrons plus tard, ce critère de sélection et de repérage des experts (et des chercheurs) tel qu'il est appliqué aujourd'hui ne constitue plus, selon nous, une garantie de compétence.

A propos de la diversité des approches, il est dit :

« la diversité dans les approches de la problématique doit être systématiquement recherchée en faisant appel à des compétences externes à l'INRA, notamment à l'étranger »

Sans contester l'intérêt qu'il y a de constituer un groupe d'experts aux compétences diverses et issus d'organismes différents, il faut observer que le recours systématique aux experts étrangers ne constitue pas nécessairement la garantie d'une plus grande qualité du travail collectif. Aujourd'hui plus encore que hier, la recherche scientifique ignore les frontières géographiques ; les collaborations scientifiques entre groupes de recherche localisés dans des pays différents sont généralement favorisées par les organismes qui financent les recherches et la grande majorité des chercheurs de qualité travaillent en collaboration avec des groupes étrangers. Dans ces conditions, le fait qu'un expert vienne d'un autre pays ne constitue pas, en soi, la garantie d'une différence de compétence ou d'approche de la question posée. Il est bon de garder en mémoire la définition de l'expert selon Mark Twain :

« an ordinary fellow from another town »

Dans la charte de l'INRA, on trouve encore à propos de l'impartialité des experts, la remarque suivante :

« les scientifiques contactés dans le cadre de l'expertise font connaître par écrit leurs liens éventuels avec certains intérêts concernés par l'expertise et susceptibles de compromettre leur neutralité »

Il est évident que la question de l'impartialité des experts est fondamentale ; il s'agit en réalité d'un problème d'honnêteté et, de ce point de vue, les scientifiques ne constituent pas un groupe particulier au sein de la société. Il y a parmi les scientifiques des gens honnêtes et des gens malhonnêtes, des gens qui se laissent acheter et des gens intègres et lorsque la décision qui pourrait résulter de l'expertise est susceptible de léser des intérêts puissants, l'impartialité des experts est particulièrement importante. Pour ne citer que deux exemples qui sont le lien entre tabagisme et cancer et le lien entre production anthropique de CO₂ et changement climatique, la puissance financière des lobbies constitués par les

producteurs de cigarettes d'une part, par l'industrie pétrolière d'autre part est telle que le risque que des experts se laissent acheter est très réel. Avec cynisme, certains disent que tout le monde peut être acheté et que la seule différence se situe au niveau du prix exigé !

La charte de l'INRA précise aussi que :

« la synthèse des avis des experts est acceptée par tous ; toute opinion divergente étant mentionnée explicitement ».

Cette remarque est, elle aussi, importante. Le groupe d'experts sollicité pour donner son avis sur une question peut fort bien ne pas aboutir à un consensus. Il n'y a rien là d'anormal mais, si telle est la situation, il est important que le rapport final fasse mention des opinions divergentes et qu'il soit accepté par tous.

Enfin, la charte de l'INRA est très explicite sur la différence de rôle entre les experts d'une part, les demandeurs d'expertise d'autre part :

« les experts ne doivent en aucun cas se substituer au demandeur d'expertise qui doit pouvoir exercer pleinement, grâce à leur éclairage, sa fonction de décideur ».

Nous avons déjà insisté sur cette séparation des rôles qui est une condition nécessaire pour que l'expertise joue pleinement son rôle. Trop souvent cette condition n'est pas remplie ; les experts ou certains d'entre eux expriment un avis sur la décision qui, à leurs yeux, doit être prise. Les médias sont friands de telles déclarations surtout si deux experts qui avaient des opinions divergentes au sein du groupe, suggèrent des décisions à prendre différentes voire opposées. Dans un tel cas, la tâche du décideur est rendue beaucoup plus difficile parce que l'opinion publique est déjà conditionnée avant que la décision ne soit prise.

Ajoutons aussi que les décideurs demandent souvent aux experts d'exprimer leur intime conviction malgré les inconnues qui demeurent. Dans les rapports d'expertise comme ceux du GIEC par exemple, les conclusions destinées aux décideurs sont plus tranchées que celles formulées dans le rapport lui-même. La stricte séparation des rôles telle qu'elle est énoncée dans la charte de l'INRA correspond donc à une situation idéale vers laquelle il faut tendre.

Une analyse de cas : la Soufrière (1976)

Si nous avons choisi de traiter des heurs et malheurs de l'expertise scientifique dans le cas de la Soufrière, c'est parce que, avec le recul, il est possible d'analyser ce cas sans passion. C'est ce qu'a fait de manière approfondie, François Beauducel, Directeur de l'Observatoire Volcanologique et Sismologique de Guadeloupe de 2001- à 2007. Son rapport de plusieurs dizaines de pages est en accès libre sur la toile (web) sous l'intitulé « Point de vue sur la Soufrière 1976 » et l'essentiel des informations factuelles reprises ci-dessous sont extraites de ce document.

La Soufrière (connue localement sous l'appellation « la vieille dame ») est un volcan situé dans le sud de Basse-Terre en Guadeloupe. C'est le seul volcan actif de l'île ; il est de type péléen et donc susceptible de donner lieu à des nuées ardentes comme celles qui ont tué 26000 personnes en 1902, sur l'île de la Martinique (éruption de la Montagne Pelée).

Dès 1975, une sismicité anormale est observée à Basse-Terre et en novembre M. Feuillard, directeur de l'Institut de Physique du Globe (IPG) de Saint Claude juge nécessaire d'alerter la préfecture. En avril 1976, M. Feuillard sollicite la venue sur place d'Haroun Tazieff, Chef du Service de volcanologie de l'IPG (Paris). H. Tazieff répond à cette invitation, prend connaissance de la situation et publie immédiatement un article dans « La Recherche » dans lequel il exprime des critiques sévères concernant les moyens humains et matériels dont dispose l'IPG de Saint-Claude.

Le 8 juillet 1976, les premières explosions phréatiques sont observées. Il s'agit là d'explosions dues à la vaporisation d'eau captive avec éjection de quantités importantes de poussières et de roches. Le phénomène est à ce point violent que 25000 personnes quittent spontanément Basse-Terre pour se réfugier à Haute-Terre. Le 13 juillet, H. Tazieff revient sur l'île avec une équipe de chimistes et un pétrographe, spécialiste des roches volcaniques (Robert Brousse). Durant le mois de juillet et le début du mois d'août, la sismicité augmente et les explosions phréatiques se multiplient. H. Tazieff a quitté la Guadeloupe pour une mission en Equateur programmée de longue date. Sur place, les scientifiques présents craignent une remontée magmatique, à leurs yeux imminente, avec le risque de formation de nuées ardentes. Alerté, le Préfet prend, le 15 août, la décision de faire évacuer la zone et c'est ainsi que tenant compte des habitants de Basse-Terre déjà partis spontanément, plus de 73000 personnes sont déplacées.

Robert Brousse émet l'hypothèse que la force de l'explosion pourrait correspondre à 1000 bombes d'Hiroshima. Haroun Tazieff qui a pu finalement être contacté à Quito cautionne l'évacuation mais conteste le risque d'une éruption magmatique cataclysmique. Revenu en Guadeloupe et après une ascension du volcan au cours de laquelle il est légèrement blessé par une chute de roches projetées par une éruption phréatique, Tazieff continue à nier la dangerosité extrême de la situation. La tension monte entre Tazieff d'une part et Brousse d'autre part.

Claude Allègre qui a été nommé Directeur de l'IPG (Paris) en juillet prend position contre Tazieff dans la querelle qui se déroule sur la place publique. Fin août, Tazieff déclare à la presse à propos de la décision du 15 août : « *Le professeur Brousse n'avait jamais vu ni étudié une éruption volcanique. Il a paniqué. Je n'aurais jamais donné de consignes d'évacuation si j'avais été à Pointe-à-Pitre* ». Rentré en France, Tazieff apprend le 6 septembre que Allègre a décidé de supprimer le service de volcanologie de l'IPG et d'assurer seul la responsabilité de toute opération volcanologique. Le 9 septembre, Tazieff dénie publiquement toute compétence volcanologique aux professeurs Brousse et Allègre et ajoute : « *la population peut sans danger travailler et habiter dans la zone évacuée après*

signature d'une décharge de responsabilité envers l'administration ». Au risque d'interrompre le récit des événements, on constate dans l'attitude de Tazieff des comportements qui, aujourd'hui, ne cadreraient pas avec les normes de l'expertise et qui, déjà à l'époque, ont interpellé le monde scientifique. Tout d'abord, Tazieff critique publiquement d'autres experts en des termes peu mesurés et, d'autre part, il se substitue aux décideurs politiques en disant ce que, selon lui, peut faire la population.

En novembre 1976, une conférence réunissant les meilleurs volcanologues au niveau mondial est organisée par le CNRS pour faire le point sur la situation sachant qu'il n'y a pas eu d'éruption cataclysmique et que la sismicité diminue. La « vieille dame » semble s'accorder un moment de repos.

Dans leurs conclusions, les participants à la conférence formulent plusieurs remarques ; ils jugent que la décision d'évacuation était justifiée mais soulignent aussi que dans le futur, il conviendra d'éviter « *une situation dans laquelle un débat public entre scientifiques et journalistes interfère avec le travail scientifique* ». Ils préconisent aussi que, dans des circonstances semblables qui viendraient à survenir, les scientifiques (nous dirions les experts) soient isolés de la presse grâce à la présence d'une personne chargée de la communication.

En conclusion de son étude du cas d'école que constitue « la Soufrière 1976 », François Beauducel examine les données de l'époque à la lumière des connaissances actuelles en volcanologie. Il conclut que, contrairement à ce que déclarait Tazieff, les risques d'une intrusion importante de magma étaient bien réels et que l'application du principe de précaution s'imposait donc. Il reconnaît aussi que plusieurs des indices utilisés par les uns pour suggérer l'évacuation et pour les autres pour en nier la nécessité étaient soit non significatifs soit erronés. F. Beauducel signale que l'éruption du Nevado del Ruiz de 1985 en Colombie qui a fait 25000 morts constitue la preuve de ce que l'absence de certitudes concernant la dangerosité d'une éruption possible ne doit pas empêcher de décider l'évacuation des zones à risques. F. Beauducel suggère même que dans le cas du Nevado del Ruiz, l'absence d'une telle décision était due, en partie du moins, à ce que certains décrivaient comme l'erreur faite en 1976, à la Soufrière. En effet, la polémique très médiatisée avait laissé des traces : l'idée prévalait, surtout dans le public, que puisque l'évacuation avait été décidée mais que l'éruption cataclysmique n'avait pas eu lieu, il n'aurait pas fallu évacuer. Il s'agit bien évidemment d'un raisonnement fallacieux mais il illustre la grande difficulté qu'éprouvent certains à faire la différence entre risque certain et risque probable. Pour éviter toute ambiguïté, il faut signaler qu'en 1985, Tazieff avait jugé que le Nevado del Ruiz présentait un grand danger pour les populations vivant à proximité.

Qu'il s'agisse de la prédiction en matière d'éruption volcanique, de tremblements de terre ou de changements climatiques, la complexité des phénomènes rend impossible, dans l'état actuel de nos connaissances, de conclure avec une absolue certitude à leur dangerosité mais cela ne doit pas conduire à l'inaction.

Les experts sont là pour éclairer les décideurs et, en matière de changement climatique, c'est ce que fait fort bien le GIEC dont le fonctionnement, comme groupe international d'experts, satisfait de manière très satisfaisante aux quatre critères précédemment énoncés. Malgré cela, certains scientifiques isolés affirment en se basant souvent sur des « preuves » très contestées que l'action de l'homme sur les changements climatiques est peu importante devant certaines causes naturelles comme les fluctuations de l'activité solaire. Agissant en tant qu'experts autoproclamés, ils en concluent que la réduction des gaz à effets de serre n'est pas une priorité. En se comportant ainsi, ces scientifiques portent une énorme responsabilité ; ils sont dans une position qui présente des analogies avec celle de Tazieff en Guadeloupe.

Tout se passe en effet comme si Allègre, pour ne citer que lui, avait repris le rôle de Tazieff ; il écrit des livres polémiques dans lesquels il tente de jeter le discrédit sur des collègues climatologues, il se répand en interviews agressives, il fustige le GIEC. Pour ceux qui connaissent l'animosité qui existait entre Tazieff et Allègre, cette reprise de rôle est fascinante à observer et serait amusante si le sujet n'était pas dramatique.

Il existe toutefois une différence importante entre le Tazieff d'hier et l'Allègre d'aujourd'hui : mis-à-part quelques membres de l'IPG de Paris proches de la direction, personne ne contestait la qualité de volcanologue de Haroun Tazieff alors que la qualité de climatologue de Claude Allègre est, elle, sérieusement mise en doute par de nombreux scientifiques et ceci tant en France qu'à l'étranger. Par ailleurs, Tazieff assumait pleinement la responsabilité de ses prises de position : affirmer qu'une éruption cataclysmique n'est pas imminente et que des dizaines de milliers de personnes peuvent regagner leurs maisons, c'est prendre le risque énorme d'être démenti par les faits le lendemain. En revanche, défendre des positions concernant le rôle de l'homme sur l'évolution du climat qui confortent ceux qui préconisent d'attendre encore avant d'agir, c'est sans grand risque immédiat pour celui qui tient un pareil discours.

Cette discussion concernant une expertise passée pour laquelle il est possible d'analyser les dysfonctionnements pose aussi et peut-être surtout la question du choix des experts. Sur quelle base faire ce choix ? Comment savoir si X a les qualités scientifiques pour être un expert ?

Comment sélectionner les experts ?

La charte de l'INRA nous suggère une réponse que nous croyons utile de rappeler :

« la compétence des experts est garantie par le mode de repérage et de sélection des experts qui s'opère sur base de leurs publications validées par leurs pairs ».

Peut-on reprocher quoi que ce soit à cette procédure qui est celle utilisée par des organismes comme le FNRS belge, le CNRS français et la très grande majorité des organismes de recherche du monde ? Personne ne contestera que les chercheurs font connaître le résultat de leurs travaux par des publications même si, dans les sciences humaines surtout, ces publications prennent souvent la forme de livres. Certaines maisons d'édition ont des comités de lecture constitués de spécialistes et donc, dans ce cas, on peut considérer que la publication est validée par des pairs mais il faut reconnaître que chez certains éditeurs, la validation se limite à fort peu de choses. Passons sur ce que d'aucuns jugeront sans grande importance puisque l'expertise dont nous traitons ici concerne prioritairement des disciplines dans lesquelles la divulgation des résultats de recherche prend la forme de publications dans des revues scientifiques, le plus souvent internationales, qui utilisent fréquemment le système des rapporteurs (referees) anonymes. Il faut toutefois nuancer cette dernière remarque ; des publications rapides, exclusivement électroniques, avec un système de relecture réduit au minimum voire non existant se multiplient ces dernières années.

Considérons le cas d'un article publié après avoir franchi une étape de validation. Est-ce pour autant que cet article est original et novateur ? La réponse est sensée venir du succès que rencontre l'article et, pour des articles scientifiques, ce succès se mesure principalement par le nombre de fois que l'article est cité. Par ailleurs, on serait tenté de considérer que le nombre d'articles publiés par un scientifique est révélateur, lui aussi, de la qualité de ce chercheur. Acceptons ce qui semble une évidence : un bon chercheur publie beaucoup de bons articles. Si par hypothèse, un bon article est un article abondamment cité on est amené à la conclusion qu'un bon chercheur est un chercheur qui publie beaucoup d'articles fréquemment cités. A ce stade, le lecteur est en droit de se demander à quoi sert d'énumérer de telles banalités. En réalité, il y a un biais dans le raisonnement qui précède et nous allons le découvrir par une analyse critique de l'usage d'un descripteur quantitatif, le facteur « H » (H index ou h index en anglais), qui a été développé pour mesurer la qualité des chercheurs sur la seule base du nombre de publications et du nombre de citations.

Le facteur « H », arme contre-productive pour choisir des experts et pour sélectionner les chercheurs

Dans tout processus de choix ou de sélection il convient d'abord de définir les critères à utiliser. Certes, pour sélectionner des chercheurs, il est facile d'affirmer « le choix doit être fondé exclusivement sur la qualité de la recherche » mais vient ensuite la question perturbante : comment mesurer la qualité ? Comment transformer un critère qualitatif en un critère quantitatif ? Est-ce qu'une publication qui fonde une nouvelle discipline ou qui révolutionne une discipline existante vaut plus ou moins que vingt articles répétitifs mais sans erreurs ?

Un physicien théoricien Jorge E. Hirsch de l'Université de Californie (San Diego), spécialiste en physique du solide et en super-conductivité, publie en 2005 un article au titre explicite : « *An index to quantify an individual research output* ». La revue choisie pour cette publication est de haut niveau puisqu'il s'agit des « Proceedings of the National Academy of Sciences » des Etats-Unis (J.E. Hirsch PNAS **102** (46) 16569-1672 (2005)).

Hirsch propose un critère simple que l'on peut décrire de la manière suivante :

un scientifique a publié, à ce jour, N articles. Si parmi ses N articles, n ont été cités au moins n fois et si les (N-n) autres articles ont été cités moins de n fois, le facteur H de ce scientifique est égal à n.

Illustrons l'application du facteur H à trois cas fictifs.

Pierre a publié 100 articles le plus souvent comme co-auteur. 7 de ces articles ont été cités 10 fois, 3 ont été cités plus de trente fois. Les 90 autres articles ont été cités moins de 10 fois. Le facteur H de Pierre est de 10.

Catherine a publié 10 articles sous son seul nom. Tous sont cités plus de 100 fois. Le facteur H de Catherine est de 10.

Georges a publié 100 articles, le plus souvent comme co-auteur. 10 d'entre eux ont été cités au moins 10 fois. Parmi ces 10 articles, 3 ont été cités pour relever des erreurs. Le facteur H de Georges est de 10

Dans son article Hirsch, effectue une analyse du facteur H de nombreux physiciens. Il en conclut que pour un physicien, un facteur H de 18 traduit un niveau compatible avec une nomination comme « full professor » et qu'un facteur H de 45 devrait ouvrir les portes de l'Académie des sciences des Etats-Unis. On constate donc que Hirsch envisage explicitement l'application du facteur H pour sélectionner des chercheurs. Etre ou ne pas être nommé « full professor » parce que son facteur H est de 19 ou de 15 : voilà la question ! Malheureusement, cette question n'est pas fictive. Aujourd'hui et dans de nombreux pays, la carrière scientifique se joue sur la valeur du facteur H de la personne soumise à sélection.

Hirsch reconnaît que les usages en matière de publications varient de discipline à discipline et que donc les seuils requis devraient être, eux aussi, dépendants des disciplines considérées. Bien évidemment, cela ne change rien au commentaire qui précède.

Hirsch envisage le cas d'un auteur dont le facteur H est relativement bas mais qui a publié quelques articles novateurs (seminal papers) avec des nombres de citations exceptionnellement élevés. Il reconnaît que dans un tel cas, le facteur H ne reflète pas pleinement la qualité scientifique de l'auteur. Hirsch reconnaît donc implicitement que le

facteur H est un outil de nivellement par le bas qui ne permet pas de déceler la véritable originalité.

Hirsch envisage aussi le cas d'un chercheur dont le facteur H élevé est obtenu grâce à des articles avec de nombreux co-auteurs. Il reconnaît que dans un tel cas, le facteur H conduit à surestimer les qualités du chercheur. L'inflation du nombre d'auteurs par article qui s'observe depuis de nombreuses années est la conséquence de l'usage qui, avant Hirsch, était fait du « citation index ». Cette inflation ne peut que s'accélérer avec la généralisation de l'emploi du facteur H.

Examinons maintenant pourquoi le facteur H ne peut servir « to quantify an individual research output » comme le dit Hirsch et que nous résumerons, de manière volontairement caricaturale, par « la quantification de la qualité d'un chercheur ». Les exemples de Pierre, Catherine et Georges cités plus haut font déjà apparaître des biais évidents de toute sélection, de tout classement d'individus basés sur la valeur de leur facteur H. D'autres remarques critiques méritent d'être formulées.

Qui oserait affirmer qu'un article cité est nécessairement un article qui a été lu et apprécié ? Si dans la publication A figurent des valeurs numériques tirées de la publication B avec référence explicite à cette publication, comment être certain que les auteurs de l'article C qui, eux aussi, font usage de ces mêmes valeurs numériques ont bien consulté la publication B ? Peut-être n'ont-ils lu que la publication A et ont-ils simplement pris note de la référence à l'article B ? Selon Ségalat¹, seuls 20% des articles cités sont lus !

Comme dans le cas fictif précédemment décrit, un article peut être cité pour être critiqué et le physicien français Anatole Abragam raconte volontiers que l'un de ses articles les plus cités doit cette caractéristique au fait qu'il contient une erreur. Plusieurs physiciens ont été trop heureux de détecter une erreur dans un article de quelqu'un qui est considéré comme un maître dans sa discipline (la résonance nucléaire). Pour l'anecdote, l'erreur avait été détectée par Abragam lui-même qui avait, en temps utile, publié une note corrective (que personne n'avait lu !).

Vient maintenant, un problème plus sérieux encore. Comment connaître le facteur H d'un scientifique particulier ? La méthode manuelle et fastidieuse consiste à déterminer pour chacune des publications le nombre de citations en utilisant pour ce faire un indice de citations. Comment savoir si cet indice est exhaustif tout particulièrement pour des articles publiés avant l'invention des publications par voie informatique ? Toutes les publications antérieures ont-elles été numérisées ? Les réponses à ces questions sont évidentes : non, l'indice n'est pas exhaustif, non, tous les articles « anciens » n'ont pas été numérisés et ceci est particulièrement vrai pour les articles publiés dans des langues autres que l'anglais. Pour

¹ Laurent Ségalat. La science à bout de souffle ? Le Seuil (Paris, 2009)

illustrer cette affirmation, prenons le cas de Vlado Prelog (1906-1998), prix Nobel de chimie en 1973 et Professeur à l'ETH de Zurich. Une grande partie de ses publications est en langue allemande et l'essentiel de son activité scientifique s'est déroulée avant l'avènement des publications par voie informatique. Le facteur H de Prelog, calculé via Google Scholar, est de 3 ! On peut évidemment considérer que ceci est sans importance puisque Prelog ne participera plus jamais à une procédure de sélection mais cela illustre combien la valeur d'un facteur H est dépendante de l'exhaustivité et de la qualité de la banque de données au départ de laquelle le facteur est calculé.

Il existe essentiellement trois banques de données bibliographiques. Le « Web of Science » de Thompson Reuters, Scopus de Elsevier et enfin Google Scholar qui présente « l'avantage » d'offrir un logiciel gratuit qui permet, sur simple demande, d'obtenir le facteur H de tout chercheur. Une fois introduit le nom du chercheur dans la case ad hoc, il suffit de lancer le programme et en quelques secondes, la valeur du facteur H apparaît. Il présente le caractère « objectif » d'un résultat fourni par l'ordinateur, instrument qui, chacun le sait, n'a pas d'état d'âme.

Si donc on est confronté à la nécessité de faire un choix entre divers chercheurs, pour savoir qui doit être nommé au sein d'une institution de recherche, ou qui doit être sélectionné comme expert, rien de plus simple que de comparer les facteurs H ; le choix peut être établi par un ordinateur et la tentation est grande de conclure que Jean avec son facteur H de 15 doit être préféré à Christiane avec son facteur H de 10. Tout membre d'une commission de classement sait que ce type d'argument est fréquemment utilisé. Comme pour s'excuser de faire appel à un facteur H auquel il ne croit pas vraiment, celui qui l'utilise ajoute généralement, un commentaire lénifiant du type : « certes je connais les critiques qui peuvent être formulées à l'encontre de ce critère mais il s'agit malgré tout de données quantitatives et objectives ». Le mal est fait : quoi qu'il soit dit après au sein de la commission, Nathalie devra porter son facteur H comme un fardeau.

Dans la littérature, on trouve des comparaisons de facteurs H de mêmes chercheurs obtenus en prenant en compte les données biographiques fournies dans des banques de données différentes². Dans cet article, écrit par une spécialiste de la bibliométrie, on trouve une analyse des avantages et inconvénients des trois banques de données précédemment citées et on arrive à la conclusion qu'aucune d'entre elles n'est exempte de graves défauts susceptibles de biaiser sérieusement le « citation index » d'un chercheur, quel que soit son domaine d'activité. Toutefois, l'intérêt majeur de l'article de Judit Bar-Ilan réside dans la comparaison qu'elle effectue entre le facteur H de 40 chercheurs israéliens abondamment cités entre 1996 et 2006 dont trois prix Nobel (un d'économie et deux de chimie). De manière assez générale, les valeurs calculées sur base de Scopus et de WoS sont voisines

² J. Bar-Ilan. *Scientometrics* (2008). **74**, 257-271. Which h-index ? A comparison of WoS, Scopus and Google Scholar

mais ce n'est pas le cas si l'on compare des facteurs H obtenus sur base de WoS ou Scopus d'une part, de Google Scholar d'autre part. Dans ce cas, les différences peuvent être importantes et, chose plus grave encore, elles varient dans un sens ou dans l'autre selon les disciplines. Les données qui suivent sont tirées du tableau 1 de l'article de Bar-Ilan. En face de chaque nom (avec la discipline entre parenthèse), on trouve le facteur H calculé selon Web of Science, Scopus et Google Scholar

Nom du chercheur	WoS	Scopus	GS
Alexander (physicien)	32	30	20
Alon (mathématicien)	14	17	27
Ciechanover (biologiste-biochimiste)	33	34	30
Cohen (immunologiste)	29	32	26
Duchovni (physicien)	32	29	15
Goldreich (computer science)	12	14	32
Harel (computer science)	9	10	22
Hershko (chimiste-biochimiste)	21	21	21
Mikenberg (physicien)	31	10	4
Netzger (space science)	28	28	18
Peleg (computer science)	8	11	21
Piran (space science)	32	32	33
Widgeron (computer science)	8	13	23

Il est manifeste que si l'on est physicien, il vaut mieux voir son facteur H calculé sur base de WoS ou Scopus plutôt que sur base de Google Scholar alors que c'est l'inverse si l'on est mathématicien ou spécialiste en informatique. Pour des chimistes et biologistes, le choix de la base semble peu important. Dans le domaine des « Space Science », les choses sont plus complexes : le choix est sans importance pour Piran mais pas pour Netzger qui lui doit, sans hésiter, refuser que l'on utilise Google Scholar pour calculer son facteur H !

Le facteur H et plus généralement tout descripteur fondé sur le nombre de publications et le taux de citation présente des effets pervers insignes qui ont été illustrés dans les pages qui précèdent. Malheureusement, les chercheurs savent qu'ils seront jugés sur base de tels

descripteurs. Ils en ont tiré des règles de conduite implicites mais que nous allons expliciter pour en faire apparaître les conséquences néfastes.

Pour avoir de nombreuses publications citées de nombreuses fois, la première condition à remplir consiste bien évidemment à publier de nombreux articles. Pour atteindre cet objectif il faut :

- 1) *Pratiquer le « saucissonnage » en fragmentant au maximum les résultats d'une recherche pouvant conduire à publication*
- 2) *Ne pas hésiter à publier plusieurs fois les mêmes résultats en prenant soin de le faire dans des journaux différents, sous des titres différents et en modifiant la forme*
- 3) *Accepter de mettre son nom sur toute publication même si le travail fourni est peu important voire nul. Il suffit de « renvoyer l'ascenseur » au plus vite*
- 4) *Accepter d'être rapporteur pour des articles de sa discipline en se limitant à la lecture de la biographie pour pouvoir suggérer d'insérer des références « manquantes », tout particulièrement celles dont on est auteur ou co-auteur*
- 5) *Engager un traducteur afin d'avoir accès à des articles de plus de trente ans publiés dans des langues autres que l'anglais pour y trouver des idées voire des résultats exploitables rapidement sans obligation aucune de citer ses sources*

Pour atteindre cet objectif, il ne faut pas perdre son temps dans des activités « inutiles » parmi lesquelles :

- 1) *Procéder à des vérifications en dupliquant des expériences ou des calculs numériques*
- 2) *S'informer à propos de domaines qui ne sont pas strictement dans sa spécialité. Ceci implique que jamais il ne faut changer de domaine de recherche*
- 3) *Lire attentivement l'article dont on est rapporteur pour suggérer éventuellement des améliorations aux auteurs*
- 4) *Ecrire des livres puisqu'ils ne sont pas repris systématiquement dans les « citations index » et surtout parce que, au mieux, ils correspondront à une publication*
- 5) *Accepter des fonctions sans doute utiles pour la communauté scientifique voire pour la société mais trop consommatrices de temps. Parmi celles-ci, accepter des tâches d'enseignement et **la participation à un groupe d'expert***

Même si les « conseils » qui précèdent paraîtront excessifs voire caricaturaux à certains, ils ne sont que la conséquence logique d'un système absurde de classement des chercheurs. La qualité d'un chercheur ne pourra jamais être décrite par un chiffre ; elle ne pourra jamais se mesurer par un nombre de publications citées.

Est-ce pour autant qu'il soit devenu impossible de déterminer qui est un bon chercheur, qui possède les qualités que l'on attend d'un expert ? La réponse à cette question est bien évidemment négative. Il existe des moyens simples d'améliorer la situation.

A titre d'exemple qui mériterait certainement d'être affiné, il faudrait demander à la personne soumise à jugement ou sélection de décrire en deux pages maximum ce qui, à ses yeux, constitue ses apports les plus importants et aussi quelles sont les marques de reconnaissance de la qualité de sa recherche auxquelles elle a été le plus sensible (invitation à présenter oralement ses résultats lors de congrès internationaux, invitations comme professeur visiteur, distinctions scientifiques,...). A ce court rapport serait joint un petit nombre de publications (au maximum cinq) jugées par le candidat lui-même comme les meilleures, les plus représentatives de son travail.

L'examen de ce document devrait être fait par des pairs qui auraient pour obligation de lire attentivement le rapport et les publications et de formuler leur avis par écrit en le justifiant. Ceci implique évidemment que les pairs eux-mêmes disposent du temps nécessaire pour faire sérieusement cette analyse et donc qu'ils n'aient plus comme unique souci de faire croître leur facteur H personnel!

On voit donc bien que toute la procédure de jugement par les pairs devrait être revue pour sortir de ce qui aujourd'hui est devenu une spirale infernale et d'autant plus infernale que ce sont les chercheurs eux-mêmes qui ont mis au point un système absurde.

Les effets pervers du mode de sélection des chercheurs sur base du facteur H a déjà donné lieu à de nombreuses publications notamment sur la toile. Parmi celle-ci, le texte de Grégoire Chamayou mis en ligne en février 2009, est à la fois lucide et percutant³. Dans un autre style, plus académique mais tout aussi convaincant, citons l'article publié dans une revue scientifique par Antoinette Molinié et Geoffrey Bodenhausen et dont le titre à lui seul est éloquent : « La bibliométrie comme arme de citation massive »⁴. L'originalité de cet article, écrit pour partie en anglais et pour partie en français, tient entre autre à ce que les méfaits de la bibliométrie sont clairement démontrés tant en physique qu'en sciences humaines. La critique des revues généralistes à haut facteur d'impact comme Science est juste et dure à la fois ; les auteurs n'hésitent pas à citer l'un de leur collègue qui a déclaré « Magazines such as Science are fit to be left on the coffee table in the common room along with Time Magazine, l'Hebdo, Spiegel and perhaps Gala ». Pour Molinié et Bodenhausen, et nous partageons ce jugement, plusieurs articles publiés dans Science « are as muddled in their argumentation as spectacular in their claims » et ne seraient jamais acceptés comme tels dans une revue spécialisée.

³ www.contretemps.eu/interventions/petits-conseils

⁴ Antoinette Molinié et Geoffroy Bodenhausen. CHIMIA 2010, **64**, 78-89

Enfin, Laurent Ségalat¹ dans son ouvrage au titre cinglant « La science à bout de souffle ? » montre lui aussi les dysfonctionnements du système actuel de sélection des chercheurs (et nous ajouterons des experts). Il met en évidence l'absurdité d'un système qui pousse les chercheurs à publier toujours plus, au point que la littérature scientifique est inondée d'articles sans intérêt aucun que Ségalat compare aux actifs toxiques qui minent la finance mondiale. Plus grave encore, Ségalat doute que le système puisse s'autoréguler.

L'expertise scientifique en grand danger

Sur base de ce qui vient d'être dit sur le dysfonctionnement majeur de la procédure de jugement par les pairs fondé principalement sur la bibliométrie, on doit donc très sérieusement s'inquiéter de la validité du choix des experts basée sur une telle procédure et cette conclusion est préoccupante parce que le fonctionnement de nos sociétés démocratiques repose en grande partie sur la crédibilité des experts et donc de l'expertise, cette expertise sur laquelle se fondent les décideurs pour prendre des décisions qui affectent notre vie personnelle et le fonctionnement de la société.

Malgré les défauts actuels du système, un jugement porté par des pairs reste sans doute la seule méthode possible de sélection des scientifiques et donc des experts mais la procédure de jugement doit être revue de manière fondamentale. Avant toute autre chose, il faudrait que les pairs qui participent à une commission de sélection de chercheurs (et) ou d'experts soient convaincus de l'extrême importance de leur tâche et qu'ils soient conscients que, ce faisant, ils contribuent à une œuvre commune. Dans le cas de la sélection de chercheurs, cette œuvre commune a pour nom « recherche de la connaissance » ; dans le cas de la sélection des experts, cette œuvre commune a pour nom « recherche des moyens à mettre en œuvre pour améliorer la vie des individus et le fonctionnement de la société ». Pour que le jugement par les pairs retrouvent sa pleine efficacité et osons le mot, sa grandeur, il faut que notre société de compétition individuelle retrouve l'altruisme comme valeur première. La phrase suivante tirée de l'ouvrage de Philippe Kourilsky⁵ « Le temps de l'altruisme » résume notre propos : « *Pour nous, le choix est clair et incontournable. Il n'y a pas d'alternative à la promotion de l'altruisme et à son inscription dans les systèmes économiques et sociaux* ».

Aujourd'hui, on assiste à un phénomène inquiétant : de manière plus ou moins diffuse, les citoyens ont perdu confiance dans le jugement des experts. De surcroît, des scientifiques qui, pour des raisons diverses, ont acquis une certaine notoriété dans le public s'autoproclament experts et peuvent, par médias interposés, critiquer le travail de ceux que l'on considère encore comme de vrais experts mais dont, bientôt, on pourra mettre en doute les qualités scientifiques compte tenu de leur mode de sélection. Les défauts majeurs

⁵ Philippe Kourilsky. Le temps de l'altruisme. Odile Jacob (Paris, 2009)

du « peer review » dans sa forme actuelle (illustrée par le recours au facteur H, comme déviance extrême) porte en germe la mort de l'expertise scientifique et, en conséquence, l'ébranlement grave de l'un des fondements des sociétés démocratiques.

Le texte de Philippe de Woot qui fait suite au présent article traite précisément de ce lien entre expertise et démocratie et des conditions à remplir pour que ce lien se renforce plutôt que de se rompre définitivement.

Mots-clé : expertise scientifique, facteur H, H index, sélection des chercheurs