

Recherche et dogmatisme : de l'improductivité du productivisme

Enka Blanchard^{*1,2,3} and Zacharie Boubli⁴

¹ Université Polytechnique Hauts-de-France, CNRS, UMR 8201 - LAMIH, F-59313 Valenciennes,

² Chaire d'Intelligence Spatiale, Université Polytechnique Hauts-de-France

³ Centre Internet et Société, UPR 2000

⁴ Conservatoire National des Arts et Métiers

24 septembre 2022

Résumé

Le renforcement du discours de l'urgence dans nos sociétés accroît la mise sous tension de la recherche. Elle fait l'objet d'attentes croissantes, parmi lesquelles se distingue une demande de productivité. Celle-ci postule d'une part la mesurabilité de la productivité scientifique et d'autre part la capacité des politiques publiques à moduler celle-ci de manière efficace et efficiente. Nous proposons une critique de ces postulats fondée sur des études empiriques provenant de multiples domaines et analysant l'organisation de la recherche dans des champs variés. Le constat d'une efficacité douteuse des politiques de recherche contemporaines nous amène à proposer l'explication d'un fondement culturel et idéologique à cet état de fait, tout en offrant des pistes pour en sortir.

Mots-clés : Productivisme, Politique scientifique, Financements compétitifs, Dogmatisme, Évaluation empirique, Urgence

1 De l'urgence de ne plus aller dans la mauvaise direction

Comme le fait remarquer Lionel Obadia dans ce même dossier (Obadia, 2021), la perception de l'urgence est facilement créatrice d'œillères sociales. Le phénomène est similaire à la *vision en tunnel* créée par les situations de stress extrême : on focalise alors l'intellect sur la résolution du problème avec une réflexion conjoncturelle et non structurelle. Le risque est alors de s'attaquer aux symptômes plutôt qu'aux causes durables. Par technophilie, les sociétés contemporaines attendent de nos jours que la recherche produise les outils de résolution de l'urgence, sans remettre en cause nos modèles sociaux. L'exacerbation des attentes envers la recherche se reconnaît dans la promptitude de certains à croire aux vertus de traitements inefficaces et non vérifiés scientifiquement, mais qui présentaient l'avantage d'offrir une solution immédiate et accessible (Mondello, 2022). Avec d'un côté une valorisation des pseudosciences — autour de figures (pseudo)-scientifiques capable de dire crédiblement ce que certains veulent entendre¹ — et de l'autre des scandales méthodologiques frappant jusqu'aux instituts prestigieux (comme l'affaire Voinnet, voir Larregue et Saint-Martin (2019)), la défiance envers les enseignants-chercheurs et chercheurs (EC) ne peut que croître. Cette défiance devient soupçon de vanité personnelle ou de biais idéologique et peut être instrumentalisée par le politique — comme avec la purge du supposé « islamo-gauchisme » souhaitée par plusieurs ministres malgré l'incohérence du concept comme élément structurant dans l'université française². Les rapports entre recherche et société

*L'auteur et l'autrice sont en ordre alphabétique, ayant contribué à parts égales.

1. Le domaine du soin et de la guérison se prête particulièrement aux croyances fondées sur des légitimités hybrides, entre charisme et rationalité détournée (Dericquebourg, 2001).

2. <https://www.cnrs.fr/fr/1-islamogauchisme-nest-pas-une-realite-scientifique>

sont ainsi tendus par l'importance des enjeux socio-techniques et les soupçons dont les EC peuvent faire l'objet. La justification de l'existence et des méthodes de la recherche peut devenir dans certains domaines (particulièrement les SHS mais aussi en médecine, en urbanisme...) un combat permanent qui se conjugue à la question de l'organisation sociopolitique des systèmes de recherche. L'enjeu de l'intégrité scientifique est revenu au centre des débats et des mesures ont été prises en ce sens, telle que l'introduction en 2022 d'un serment d'intégrité scientifique pour les docteurs³.

Au-delà de la question du rapport du public à la communauté scientifique, un enjeu critique concerne l'allocation des moyens de la recherche — publique comme privée. Le désengagement tendanciel de la puissance publique dans les dernières décennies met toujours plus en avant le souci d'une optimisation et d'une rationalisation du financement de la recherche (Chaudret, 2013) sous l'effet bien connu du *New Public Management*⁴ (Park & Berry, 2014). Pour cela ont été créés des instruments d'évaluation toujours plus raffinés visant à évaluer la productivité des EC pour s'assurer que les pouvoirs publics ne soutiennent que les « meilleures recherches » (Aghion & Cohen, 2004). Le financement de la recherche publique s'est donc orienté vers la rationalisation justifiante de l'élitisme, que l'on peut définir ici comme une répartition hautement inégale des moyens, concentrés vers une faible proportion de la population considérée au nom d'un mérite supposé. Cela amène d'ailleurs le débat politique sur la recherche à se concentrer sur la rationalité de l'allocation des ressources au détriment de leur augmentation, qui se manifeste notamment par la recherche permanente d'indicateurs de performance (Muller, 2017). Il en a résulté une mutation de la carrière-type des EC, reflétée dans un fonctionnement par projets mais surtout une compétition croissante pour les emplois d'EC statutaires, dont la raréfaction n'est que partiellement compensée par la création de postes intermédiaires — comme les chaires de professeur junior. Le modèle dit américain, qui tend à s'imposer comme modèle global, repose sur une compétition exacerbée basée sur le « publish or perish », récompensant les EC plus performants au détriment des EC ordinaires et de la qualité de vie de l'ensemble du personnel de la recherche. Un modèle français fondamentalement différent a temporairement existé, privilégiant la sécurité de l'emploi et l'équilibre de vie à la rémunération (Chamayou, 2009). On ne peut douter de la performance de ce modèle alternatif, notamment en mathématiques où l'attractivité du système lui a permis de briller au niveau de pays ayant des populations et des budgets plusieurs fois supérieurs — selon les normes d'évaluation aujourd'hui en vigueur (comme le nombre de médailles Fields). Il semble donc naturel d'interroger les justifications apportées à l'évolution vers un modèle plus compétitif — d'autant plus que les rigueurs budgétaires françaises limiteront l'attractivité du métier de chercheur en France par rapport à nos voisins (si la qualité de vie au travail est équivalente).

Or, un examen poussé des formes de la rationalisation des politiques publiques montre que celles-ci ne s'applique pas systématiquement à elle-même et que les optimisations résultantes peuvent être moindres que les coûts liés à l'optimisation⁵. Il nous semble donc propice de faire un point sur l'évaluation de la recherche, une évaluation de l'évaluation, suivant les propos⁶ de Pierre Azoulay (2012) : « Il serait une étrange coïncidence que les systèmes qui ont fonctionné au XX^e siècle fonctionnent aussi bien au XXI^e siècle. Nous prendre nous-mêmes comme objets d'étude pourrait bien exposer certains défauts de la communauté scientifique et nous exposer à la critique de politiciens à l'affût de raisons de réduire les financements scientifiques. Mais la seule alternative serait la sclérose progressive des institutions scientifiques les plus respectées. »

Cet article propose donc une analyse de trois piliers sur lesquels reposent les politiques publiques

3. <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000046228965>

4. Des effets similaires s'observent dans d'autres domaines, notamment celui de la culture (Arfaoui, 2019).

5. Un exemple notable est la politique d'optimisation menée par le Département du travail et des retraites (DWP) du Royaume-Uni, qui voulait réduire le nombre de bénéficiaires de prestations sociales pour le handicap. La méthode passait par des entretiens individuels externalisés auprès de plusieurs cabinets de conseils (principalement Atos au départ). Les critiques à l'époque croissantes sur les milliers de morts évitables et les coûts sociaux causés par ces politiques n'avaient pas eu d'impact. Mais l'intéressant est la réaction face à la sous-performance des cabinets : taux d'erreurs de plus de 30% au lieu des 3% annoncés, vitesse de traitement des dossiers deux fois inférieure à la prévision, accumulation de nombreux retards. Cette sous-performance et les coûts démesurés qu'elle entraînait (en dizaines de millions de livres annuels pour les seuls frais de justice, en centaines de millions de livres en frais supplémentaires facturés par les cabinets) auraient dû, dans une perspective d'optimisation budgétaire, mener à l'annulation du contrat dès la première année, ce qui fut loin d'être le cas (Stewart, 2018).

6. Traduction libre de l'autrice et l'auteur.

de la recherche : la mesurabilité de la productivité des EC, l'efficacité des politiques publiques visant à augmenter la productivité et enfin l'efficience de celles-ci étant donnés les coûts d'évaluation directs comme indirects. Nous nous basons sur plusieurs décennies de travaux principalement anglophones afin de donner un état de l'art sur la productivité scientifique et le productivisme des politiques publiques afférentes. Cela nous amène à formuler l'hypothèse que ces dernières sont fortement influencées par un dogmatisme organisationnel. Celui-ci pourrait bel et bien réaliser une intuition de Leo Szilard en 1948 sur comment ralentir le progrès scientifique⁷ : « mettre en place une fondation dotée d'un budget conséquent.[...] Ainsi, les meilleurs scientifiques seraient occupés en commission d'examen des demandes, et les chercheurs en besoin de fonds se concentreraient sur les sujets apparaissant les plus prometteurs pour des résultats publiables. » Tout porte donc à croire que l'urgence de la recherche est de sortir d'une vision tunnel dogmatique et d'appliquer les connaissances scientifiques contemporaines à l'élaboration des politiques publiques.

2 La bonne direction est-elle celle du productivisme ?

On peut supposer que la plupart des États finançant des travaux de recherches visent à optimiser à la fois la qualité et la quantité de ces recherches — l'interprétation de ces termes étant sujette à caution — pour un investissement donné en moyens humains, techniques et budgétaires⁸. Notre but ici n'est pas de remettre en question ce modèle mais plutôt de vérifier s'il atteint les objectifs qu'il se fixe. Plus explicitement : les politiques publiques de gestion de la recherche visant à augmenter la productivité suivent-elles des méthodes appropriées pour atteindre leurs cibles ? Une réponse négative pourrait être causée par plusieurs facteurs : des luttes de pouvoir, des biais idéologiques, une réflexion insuffisamment poussée sur les effets de ces politiques, ou simplement un manque d'information. Ces oeillères sociales ne seraient que renforcées par un discours d'urgence permanente ne permettant pas la réflexivité et l'observation des conséquences sur le temps long.

- Nous proposons d'analyser trois postulats sur lesquels semblent reposer ces politiques publiques :
- il est possible de mesurer avec une précision suffisante la productivité des EC, tant au niveau collectif qu'individuel ;
 - il est possible d'augmenter la productivité de l'écosystème de la recherche par différents mécanismes à la fois de sélection (recrutements par concours, financements des projets par compétition, concentration des moyens sur les personnes/laboratoires aux performances les plus élevées) et d'incitation (primes d'excellence, avantages financiers ou sociaux) ;
 - les coûts d'évaluation et de mise en œuvre de ces politiques publiques sont plus que compensés par les gains en productivité, y compris en intégrant les potentiels effets délétères.

Les trois sous-sections suivantes examinent la vraisemblance de ces postulats un par un.

2.1 La productivité scientifique est-elle mesurable ?

Le premier postulat établit que la production et la productivité des EC — que l'on définit ici comme la production rapportée aux ressources, notamment temporelles — est mesurable (avec une certaine précision). Cette mesurabilité est au coeur de nombreux contentieux depuis plusieurs décennies, avec au moins trois questions centrales :

- Quels sont les objets (*productions scientifiques*) et sujets (EC ou groupe de recherche) devant être évalués ?
- Quelles méthodes l'évaluation doit-elle suivre : moyens qualitatifs (à priori par un jury de pairs) ou quantitatifs (le plus souvent par des outils bibliométriques et plus généralement scientométriques) ?
- Quelles échelles spatiales et temporelles l'évaluation doit-elle considérer ?

7. Traduction libre, source : <https://library.ucsd.edu/dc/object/bb33804055>

8. Cet investissement n'est pas toujours fixé à l'avance — par exemple avec les appels *Oh Risque* de l'Agence nationale de la recherche (ANR)— et on peut d'ailleurs s'attendre à des rendements décroissants à partir d'un certain seuil.

Une mesure de bonne qualité doit donner des réponses à ces questions, sans négliger de considérer l'impact de la mesure sur ce qui est mesuré (nous y reviendrons). Surtout, les évaluations ne mesurent que le passé, en supposant qu'une haute productivité récente indique nécessairement une haute productivité future. Cela ignore la variabilité inhérente de la recherche et augmente la pression à exceller en début de carrière en faisant de la recherche en urgence pour optimiser les indicateurs.

Ces *productions scientifiques* devant être évaluées sont souvent classées en trois catégories :

- les missions d'enseignement, d'encadrement, d'administration et de diffusion/vulgarisation ;
- les publications : revues, ouvrages ou chapitres d'ouvrages, actes de conférences ;
- les aspects économiques, à la fois les valorisations (brevets, collaborations industrielles) et la capacité à obtenir des financements : bourses, projets compétitifs type ANR ou ERC (*European Research Council*).

En dépit d'efforts du corps universitaire pour souligner les liens entre recherche et enseignement, ce dernier reste le plus souvent vu comme une corvée (ou « charge ») par les EC et administrateurs. L'effet est donc de l'occulter au profit des deux autres, qui forment l'objet central des mesures de productivité (Reymert, Jungblut, & Borlaug, 2021).

Commençons par l'évaluation des publications, tant quantitative que qualitative. Dix ans après la déclaration de San Francisco (American Society for Cell Biology, 2012), les problèmes soulevés à l'époque restent d'actualité, ayant mené en 2022 à l'élaboration d'une convention européenne sur la réforme de l'évaluation⁹, dont les effets restent à voir. D'ici là, les évaluations d'EC comme de laboratoires se basent toujours sur les outils bibliométriques. Ceux-ci reposent essentiellement sur le nombre de publications (en particulier dans les revues « prestigieuses »¹⁰) et les citations — utilisées comme indicateur de « qualité » des publications à la fois dans l'évaluation et comme norme dans de nombreuses études bibliométriques (McKiernan et al., 2019). La bibliométrie sert non seulement de filtre (Reymert, 2021) mais aussi d'outil déterminant dans le recrutement et la promotion des EC (McKiernan et al., 2019 ; Van Dijk, Manor, & Carey, 2014), l'Italie¹¹ offrant un exemple européen emblématique (Demetrescu et al., 2020). L'utilisation de la bibliométrie reste pourtant considérée comme un fait accompli, seule garante de la rigueur des évaluations face au flou du qualitatif (Muller, 2017). Les éléments de preuve empiriques qu'on pourrait attendre à l'appui du bibliométrisme se retrouvent moins souvent dans le discours de ses promoteurs que de ses critiques (Chawla, 2021).

Or ces mêmes critiques ne remettent que rarement en question la notion même d'évaluation, pour laquelle ils estiment le plus souvent que des évaluations qualitatives par des jurys d'experts sont plus satisfaisantes. L'exemple central de cette logique réside dans le recrutement compétitif des EC. Étant donnée l'augmentation des candidatures, il n'est pas réaliste de lire attentivement à la fois les projets et toutes les publications des candidats. Pour hiérarchiser ceux-ci, on se reporte alors aux jugements déjà effectués par les pairs — par exemple les évaluations d'articles par les comités de revue. Il émerge cependant une critique de la fiabilité de ces signaux (Garcia-Costa, Squazzoni, Mehmani, & Grimaldo, 2022), notamment grâce aux données empiriques de l'étude *NeurIPS* de 2014 (Cortes & Lawrence, 2021). Lors de cette expérience effectuée sur l'une des plus prestigieuses conférences¹² d'intelligence artificielle, deux comités parallèles évaluaient 1 678 articles, dont 170 articles en commun — sans savoir quels étaient ces derniers. Ces deux comités ne tombèrent d'accord que sur 51% des articles acceptés (là où le hasard aurait donné 23.5%). Chaque article avait une note attribuée entre zéro et dix par au moins trois évaluateurs, avec une corrélation de 0.55 entre les scores moyens des deux comités, corroborant les études précédentes sur l'importance de la composante aléatoire dans l'évaluation (Bornmann, Mutz, & Daniel, 2010). En outre, parmi les articles acceptés, aucune corrélation n'existait entre la note d'un

9. <https://www.scienceeurope.org/our-resources/agreement-reforming-research-assessment/>

10. Le *facteur d'impact*, créé au départ par des bibliothécaires pour déterminer quelles revues obtenir en priorité, a été détourné pour en faire une échelle de qualité des revues (Stephan, Veugelers, & Wang, 2017). Le classement des revues a glissé vers un classement des EC qui y publient — malgré la facilité qu'il y a fausser les indicateurs individuels, par exemple par l'autocitation (Biagioli & Lippman, 2020 ; Chawla, 2018).

11. L'Agence nationale d'habilitation italienne exigeait avant tout recrutement ou promotion que les candidats aient des scores bibliométriques dépassant la médiane des personnels ayant le grade supérieur. Ces exigences furent remplacées en 2016 par des seuils calculés par un algorithme dont le fonctionnement n'est pas public (Demetrescu, Finocchi, Ribichini, & Schaerf, 2020).

12. Le domaine de l'informatique a comme particularité qu'il est souvent plus prestigieux (et difficile) de publier dans les actes des conférences de premier rang que dans les meilleurs journaux.

article et le nombre de citations reçues sept ans plus tard. Le fait de publier dans un journal à haut facteur d'impact n'est pas non plus une preuve de qualité, une étude récente sur 45 144 articles de sciences cognitives ayant non seulement montré que le nombre de citations est négativement corrélé à la rigueur statistique mais aussi que le facteur d'impact du journal est très faiblement corrélé à la rigueur statistique et négativement corrélé à la répliquabilité des études (Dougherty & Horne, 2022). Ces résultats empiriques sur l'inadéquation de l'évaluation par les pairs ne sont pas non plus nouveaux : une étude de 1998 testant 262 évaluateurs en médecine montrait déjà que la plupart des erreurs parfois grossières dans un article n'étaient pas détectées par les évaluateurs (Baxt, Waeckerle, Berlin, & Callahan, 1998). Si l'évaluation d'un simple article par un jury d'experts du domaine est si peu fiable, on ne peut que douter de la qualité du jugement de jurys moins spécialisés chargés de recruter un collègue ou d'attribuer des financements — une autre étude (Pier et al., 2018) ayant d'ailleurs montré une absence totale de corrélation dans les évaluations de projets financés par les Instituts américains de la santé (NIH). Ces résultats n'indiquent d'ailleurs pas que les évaluateurs bâclent leur travail, mais potentiellement que ce dernier est intrinsèquement sujet à de l'incertitude.

L'impact de ces jugements imprécis est hélas amplifié par un effet Matthieu — selon lequel un écart initial non justifié a tendance à augmenter avec le temps. À l'échelle de la carrière d'un EC, on souhaiterait qu'une évaluation initiale erronée n'affecte pas les suivantes. Les résultats d'évaluations devraient donc idéalement être relativement stables dans le temps, mais surtout, une évaluation ne devrait pas avoir d'impact sur les suivantes au-delà de ce qu'elle révèle de la qualité intrinsèque de l'EC. Plusieurs études indiquent cependant le contraire, un des cas les plus visibles concernant l'attribution des financements par l'Organisation pour la recherche scientifique aux Pays-Bas (Bol, de Vaan, & van de Rijt, 2018). Les candidats recevant une note et l'attribution des bourses ne dépendant que de cette note, il est possible d'étudier les trajectoires d'EC situés juste au-dessus et au-dessous du critère d'attribution. On observe alors que les financements totaux (hors bourse initiale) reçus par les lauréats sur les huit années suivantes sont entre deux et trois fois supérieurs à ceux reçus par les non retenus. Il n'y a pourtant pas de différence statistique ni au sein des « gagnants » (peu important les notes) ni au sein des « perdants »¹³. Les promotions au grade de professeur sont similairement affectées (Bol et al., 2018). Cet effet ne concerne pas uniquement les financements mais aussi les citations : un article déjà cité accumule inéluctablement d'autres citations, y compris lorsqu'il est retiré pour faute (Teixeira da Silva & Dobránszki, 2017). Une étude sur des articles dupliqués montre que lorsqu'un même article est publié dans deux journaux différents (pour des motifs variables), le nombre de citations est hautement dépendant du journal — à qualité identique de l'article (Larivière & Gingras, 2010). Cet effet se retrouve enfin dans la possibilité même de publier dans certaines revues. Bien qu'il y ait un nombre croissant de revues scientifiques de qualité, la proportion des articles publiés dont le premier auteur a déjà publié dans la revue est passée de moins de 5% dans les années 1950 à plus de 20% récemment — pour une étude concernant presque exclusivement des journaux où les auteurs ne sont pas anonymes pendant l'évaluation — menant à une concentration d'autant plus grande au sein des journaux prestigieux (Fire & Guestrin, 2019).

Étant donné le caractère arbitraire et aléatoire des évaluations, tant sur les articles que sur les financements, nous nous retrouvons confrontés à une difficulté fondamentale dans la mesure de la productivité des EC, tant sur le plan qualitatif qu'avec des outils scientométriques. Les améliorations de ces derniers, par exemple avec les *altmetrics*¹⁴, ne font d'ailleurs qu'alimenter une course aux armements sans solution stable envisageable à long terme (Griesemer, 2020).

Même en faisant l'hypothèse que des mécanismes de standardisation et de responsabilisation puissent réduire l'arbitraire de l'évaluation, la mesure de la productivité ferait encore face à un obstacle majeur : la loi de Goodhart. Dérivée de la loi de Campbell provenant de l'analyse des politiques publiques, cette loi énonce que « lorsqu'une mesure devient un objectif, elle cesse d'être une bonne mesure », ce qui se vérifie de plusieurs manières (Rodamar, 2018). Tout d'abord, l'usage d'outils scientométriques a en-

13. Une partie de cette différence peut être expliquée par une tendance à moins chercher d'autres financements en cas d'échec initial. Le facteur dominant affectant l'obtention de financements futurs semble toutefois être l'obtention d'un financement initial.

14. Les *altmetrics* sont des indicateurs alternatifs censés capturer la portée en ligne au-delà de la citation : nombre de vues, téléchargements, mentions sur les réseaux sociaux...

couragé l'évolution des modes de publication et l'augmentation drastique du nombre d'articles¹⁵ ainsi que du nombre de citations (Fire & Guestrin, 2019). Ces changements rendent difficiles non seulement les comparaisons entre domaines — le nombre médian de citations variant d'un facteur au moins 30 entre ceux-ci — mais aussi entre sous-domaines très proches (e.g. plusieurs sous-domaines d'étude de l'ARN). Ils rendent aussi presque impossibles les comparaisons avec un écart temporel, avec des écarts non négligeables à l'échelle de l'année (Patience, Patience, Blais, & Bertrand, 2017). Cette augmentation du nombre d'articles n'est pas uniquement due à une productivité plus haute ou à un nombre d'EC plus important, mais aussi à un changement des modes de publication (favorisant par exemple les articles de revues aux chapitres de livres et rapports (Sîle & Vanderstraeten, 2019)). La longueur des articles a aussi considérablement baissé, avec un nombre de pages presque divisé par deux depuis 1950, transformant l'unité de base de la mesure (Fire & Guestrin, 2019).

Au-delà des changements de stratégie causés par l'incitation à plus publier et à plus se faire citer, il ne faut pas oublier que l'importance croissante accordée aux mesures — plutôt qu'aux articles sous-jacents — facilite et encourage la fraude sous toute ses formes (Biagioli & Lippman, 2020). Alors que la crise de réplication s'éternise et montre que les erreurs sont difficiles à déceler dans de nombreux domaines, notamment les sciences naturelles (Borsboom, 2006 ; Mullard, 2021 ; Rodgers & Collings, 2021), l'augmentation de la fraude pose un problème non trivial ayant déjà des conséquences graves, par exemple avec le scandale récent sur l'hypothèse amyloïde pour la maladie d'Alzheimer (Piller, 2022).

2.2 La productivité scientifique peut-elle être augmentée par les politiques publiques ?

Si le premier postulat concernait la mesurabilité de la production scientifique, le deuxième concerne la possibilité d'augmenter cette production par des politiques publiques. On pose donc la question de l'augmentation de la productivité — considérée ici naïvement comme une augmentation des publications (ou des publications hautement citées) — généralement modulée selon deux logiques distinctes. La première est l'incitation : promettre des récompenses (notamment financières) aux EC s'ils augmentent leurs performances¹⁶. La deuxième méthode est la sélection : faire en sorte que les ressources publiques soient concentrées le plus possible sur les EC les plus productifs. Il s'agit donc non seulement de recruter les EC les plus prometteurs mais aussi de donner des moyens additionnels (bourses de recherche, financements de thèses et postdocs, matériel de pointe) aux équipes les plus performantes. Examinons d'abord les problèmes avec les mécanismes compétitifs de financement.

La première question est contre-intuitive : les mécanismes de compétition sont-ils toujours sélectifs ? On peut en effet imaginer un système où chaque financement est sélectif, mais où ces derniers sont en nombre suffisant pour que chaque candidat en trouve un — non sans effort considérable. Une étude de 2011 sur les financements fournis par les NIH a observé un effet négligeable de l'obtention de bourses compétitives (de 1,7M\$ en moyenne) sur les publications (avec une augmentation de la productivité autour de 7%, soit une publication dans les cinq années suivant l'obtention) (Jacob & Lefgren, 2011). Laissant de côté l'explication potentielle qu'une bourse n'augmente pas la productivité en publications¹⁷, l'explication alors proposée est que la non-obtention de la bourse peut se compenser par l'obtention d'autres sources de financement (principalement industriels vu le domaine). On aurait donc un exemple de la situation imaginée, où plusieurs mécanismes sélectifs forment un système qui n'est pas sélectif à l'échelle globale. Tout l'édifice compétitif serait alors superflu.

15. Comme le disait Grégoire Chamayou en 2009 : « Ne faites pas de la recherche : écrivez des papiers » (Chamayou, 2009).

16. Cette méthode se base sur une perception que de nombreux EC manquent de motivation et travaillent peu, perception énoncée notamment par l'ancien président Nicolas Sarkozy. Attaquant la basse productivité en terme de publications des EC français de certains domaines, ce dernier les accusait de ne venir au laboratoire que parce que « il y a de la lumière, c'est chauffé » (Chaudret, 2013).

17. Une critique de cette étude donne une autre interprétation, selon laquelle la différence négligeable n'est pas due à une saturation du marché des financements mais à une décorrélation plus fondamentale entre financement et production scientifique. L'explication proposée par les auteurs de l'article initial serait alors une explication ad-hoc motivée par une croyance injustifiée en l'applicabilité des mécanismes de marché sur la production universitaire (Muller, 2017).

La deuxième question concerne le lien entre productivité et élitisme des financements — c’est-à-dire l’impact d’une distribution inégale des crédits de recherche. Plusieurs études offrent des critiques de ce lien supposé. L’analyse de plusieurs pays de l’OCDE sur la période 1987-2006 a par exemple montré une absence de différence notable de productivité, à la fois entre pays dont le degré d’élitisme varie, mais aussi après l’adoption au sein d’un pays de politiques visant à renforcer l’élitisme (Auranen & Nieminen, 2010). Une deuxième étude sur 12 720 EC québécois sur la période 2000-2013 analyse l’impact des financements sur la production d’articles et d’articles hautement cités ainsi que le nombre de citations des articles produits (Mongeon, Brodeur, Beaudry, & Larivière, 2016). Le système québécois est hautement élitiste : 10% des chercheurs reçoivent entre 50% et 80% des financements nationaux selon les disciplines. En intégrant les financements provinciaux québécois, ces parts correspondent aux 20% des chercheurs mieux financés. Dans tous les domaines, les trois indicateurs montent extrêmement rapidement quand on passe de l’absence de financement à un faible financement (20-50k\$ CAD annuels, les financements étudiés allant jusqu’à 2M\$ CAD), avec une augmentation plus lente au-delà. Cependant, à part sur la recherche médicale, les indicateurs atteignent ensuite un maximum¹⁸, après quoi les indicateurs diminuent quand les financements augmentent. On observe des rendements marginaux non seulement décroissants, mais même négatifs quand on atteint des financements élevés¹⁹. De plus, les 20 EC les plus financés dans chacune des disciplines (ayant 3 000 à 5 000 EC chacune), malgré un budget 27 à 40 fois supérieur, ne publiaient qu’entre 5 et 19 fois plus²⁰, avec un impact faible voire négatif sur les citations. Ces résultats rejoignent d’autres études portant sur la productivité par dollar de financement (Lauer, Danthi, Kaltman, & Wu, 2015 ; Wahls, 2018b). Par exemple, en recherche médicale, certaines universités prestigieuses (comme Harvard, Stanford, Johns Hopkins...) ont de plus hauts taux de succès sur tous les indicateurs de financements (dollars reçus au total et par EC, proportion des projets lauréats...) (Wahls, 2018a). Cependant, les universités moins prestigieuses (comme les universités du Dakota du Nord ou du Nebraska) publiaient plus par dollar de financement (65% en moyenne) et recevaient plus de citations (35% en moyenne).

Un autre danger lié aux politiques élitistes est que l’augmentation de la productivité peut se faire aux dépens de la diversité, sur plusieurs plans. Tout d’abord, les sujets traités peuvent être transformés, certaines pistes de recherche ignorées non pas car elles semblent moins importantes mais parce qu’elles auront moins de débouchés à court terme ou seront plus difficiles à financer (Laudel, 2006). La recherche en urgence se fait ainsi aiguillon néfaste sur l’exploration scientifique à long terme. L’élitisme peut aussi renforcer des biais existants portant non pas uniquement sur les projets mais aussi sur les EC portant ces projets. Ces problèmes ont déjà été fréquemment soulignés, avec notamment des biais racistes et sexistes dans les financements qui perdurent (Madsen & Aagaard, 2020 ; Taffe & Gilpin, 2021) — mais aussi du favoritisme en faveur des EC de certaines institutions et une pénalisation des recherches interdisciplinaires (Banal-Estañol, Macho-Stadler, & Pérez-Castrillo, 2019).

Ayant explicité les problèmes principaux posés par les mécanismes de sélection, nous pouvons désormais nous pencher sur les mécanismes d’incitation, correspondant typiquement aux récompenses monétaires pour des publications prestigieuses. Le prestige mais surtout le facteur d’impact sont au coeur de politiques nationales de nombreux pays tels que la Corée du Sud et la Turquie, mais surtout la Chine (Franzoni, Scellato, & Stephan, 2011). Ces politiques y sont poussées à l’extrême, avec des primes pouvant atteindre 20 années de salaire pour un article dans *Science* ou *Nature*, qu’accompagnent bien entendu des effets délétères sur la qualité et l’intégrité scientifique (Quan, Chen, & Shu, 2017). Le modèle sous-jacent d’agent rationnel dont l’effort est corrélé à la rémunération est critiqué depuis longtemps, notamment dans le secteur public (Park & Berry, 2014). En effet, une analyse plus détaillée

18. Ce maximum est atteint autour de 80k\$ CAD pour les sciences sociales, comparé à un financement moyen de 16k\$, et à 300k\$ CAD pour les sciences naturelles, comparé à un financement moyen de 53k\$.

19. Il existe une étude notable indiquant un résultat contraire sur des données similaires au Canada (Ebadi & Schiffauerova, 2016). Cette étude montre un rapport positif entre financements, qualité et quantité des publications, ainsi qu’une qualité supérieure des productions des chercheurs industriels. Cependant, cette étude fait une impasse majeure en comparant les indicateurs bibliométriques sans normaliser entre domaines différents. Il n’est donc pas étonnant qu’on observe une différence de « qualité » — évaluée en fonction du nombre de citations — entre des chercheurs industriels en nanotechnologie et génétique d’un côté et des chercheurs universitaires en mathématiques ou en littérature.

20. Par ailleurs, cette étude ne contrôle pas que les EC publiant plus ont tendance à obtenir plus de financements, et donc que même sans ces financements, les EC correspondant auraient probablement toujours publié plus que la moyenne.

des motivations intrinsèques et extrinsèques montre que dans certains cas — notamment pour les tâches plus intellectuelles et créatives — ajouter une incitation financière (donc extrinsèque) peut nuire à la motivation intrinsèque et donc à la performance (Hennessey & Amabile, 2010). La quantité de travail fourni — sans parler de la qualité — peut donc diminuer en présence de primes à l'excellence. En l'occurrence, une étude sur les propositions d'articles à la revue *Science* entre 2000 et 2009 a montré que, en présence de primes récompensant financièrement²¹ la publication dans des revues prestigieuses, le nombre de propositions augmentait grandement, alors que le nombre de publications acceptées diminuait (de manière non significative cependant) (Franzoni et al., 2011). Cela ne prend naturellement pas en compte l'impact de ces primes sur l'incitation à publier des recherches basées sur des données faussées²².

Les problèmes sus-cités ne doivent pas masquer une transformation plus fondamentale. De longue date, les politiques publiques de la recherche visent à augmenter la productivité elle-même. Ces dernières décennies ont pourtant vu l'émergence d'un glissement téléologique dangereux. L'augmentation de la qualité et de la quantité des productions scientifiques est en voie de transformation en un fétichisme des indicateurs et des classements comme buts propres, avec par exemple l'objectif annoncé par l'exécutif français d'avoir deux établissements dans le top 20 du classement de Shanghai (Anonymous, 2010). C'est ainsi qu'apparaissent un ensemble de mesures visant à augmenter le classement artificiellement — comme le regroupement d'universités, ayant parfois d'autres motivations mais justifié par cet impératif de performance au classement (Barats, 2019). Ces mesures n'auront au mieux qu'un effet négligeable sur la production scientifique²⁴ mais ont surtout des coûts.

2.3 Les politiques publiques scientifiques sont-elles efficaces ?

Nous avons considéré dans la section précédente la possibilité que l'élitisme scientifique soit fécond. En supposant qu'il le soit un tant soit peu, il n'en découle pas automatiquement que les effets positifs dépassent les coûts engendrés par cet élitisme.

Les premiers coûts reposent directement sur les EC. Les efforts investis dans l'élaboration des demandes de financements sont principalement motivés²⁵ par l'espoir d'un succès éventuel — offrant à la fois un soutien matériel et un gain en prestige. Cependant, dans la plupart des pays à financements compétitifs, ces demandes de financements se soldent dans 80% à 90% des cas par des échecs. Le temps passé à obtenir et gérer des ressources est donc non négligeable et peut dépasser 40% du temps de travail dans certains pays comme les États-Unis (National Science Board, 2014). Par exemple, une étude de 2013 (Herbert, Barnett, Clarke, & Graves, 2013) a montré que la préparation d'une demande de financement au NHMRC (organe principal de financement de la recherche médicale en Australie) prenait en moyenne 35 jours ouvrés de travail d'EC — auquel il faut ajouter le temps de travail du personnel technique et administratif du laboratoire. Sans compter les coûts d'évaluation, la simple prise en charge des salaires pendant cette durée représentait 14% du budget alloué aux projets de recherche — alors même que le taux de succès était à l'époque de 21%, comparé à 13% en 2019 et 2020 (Australian National Health and Medical Research Council, 2021). Cette préparation empiète non seulement sur la recherche mais aussi sur le temps personnel et les vacances et est un facteur de stress majeur (Herbert, Coveney, Clarke, Graves, & Barnett, 2014).

Ces compétitions ont ensuite un coût sur le système de la recherche lui-même, notamment par le

21. La présence d'incitations moins immédiates concernant l'avancée de carrière avait cependant un effet positif sur le nombre de publications acceptées.

22. La course à la publication a été mentionnée comme facteur ayant incité à la falsification lors du rapport de contrôle de l'IHU Méditerranée infection par l'Inspection générale des affaires sociales (Desailly-Chanson, Penaud, Vilboeuf, Elshoud, & Piozin, 2022), avec au moins 86 articles concernés par des problèmes signalés sur *Pubpeer* par Elisabeth Bik. L'IHU a d'ailleurs reçu de grandes sommes via SIGAPS²³, en partie grâce aux publications prolifiques de ses membres, même si celles-ci sont concentrées dans des revues de moins bonne « qualité », tenues par les membres de l'IHU, voire même des revues prédatrices (Maisonneuve, Plaud, & Caumes, 2020).

24. Un des arguments utilisés pour justifier ces systèmes est l'attractivité à l'international, mais cela ne mène qu'à une compétition où le système entier perd quand tous les membres sont forcés de concourir.

25. En contrepoint, le fait de participer à ces processus compétitifs est parfois souligné comme pouvant avoir en soi un impact positif sur la production scientifique (Ayoubi, Pezzoni, & Visentin, 2019).

nombre d'évaluateurs experts requis, y compris quand ces derniers sont bénévoles²⁶ et y affectent du temps autrement disponible pour la recherche. Ces coûts d'évaluation sont d'autant plus élevés par rapport aux sommes en jeu que le système est élitiste (un seuil important est celui des 50% de projets non financés, seuil plus élevé qu'à peu près tous les systèmes de financement compétitif). En pratique, certains modèles montrent que les systèmes actuels fonctionnent déjà potentiellement à perte, avec un coût d'évaluation supérieur au budget total alloué (Gross & Bergstrom, 2019). Quand cette évaluation comporte une part d'aléatoire, l'augmentation de la sélectivité amène des effets pervers. L'intérêt marginal d'une dépense de temps et d'effort pour améliorer une demande de financements chute à la fois avec la sélectivité et avec l'importance du hasard. Quand le facteur chance est perçu comme prépondérant et qu'un seuil de sélectivité est franchi, la stratégie gagnante devient donc de produire un effort minimal par demande mais d'augmenter massivement le nombre de demandes (Roebber & Schultz, 2011). Cela augmente mécaniquement les coûts d'évaluation et donc l'inefficience du système (Herbert et al., 2014). Même les contraintes comme celles présentes pour les bourses du Conseil européen de la recherche (ERC) — où une mauvaise évaluation peut mener à l'interdiction de demander une bourse pendant plusieurs années — ne peuvent suffire à compenser pleinement ces effets.

Face à ces coûts élevés, des modèles théoriquement plus efficaces ont été proposés (Gross & Bergstrom, 2019). Une première option est la simplification des démarches afin d'avoir des demandes de financements beaucoup plus courtes, n'incluant par exemple que le contenu scientifique et aucune composante administrative (Herbert, Graves, Clarke, & Barnett, 2015). Cette solution va cependant à l'encontre de la demande publique de transparence et de contrôle rigoureux des dépenses publiques tout en se heurtant à l'inertie des administrations face aux simplifications majeures (National Science Board, 2014). De plus, l'effort lié aux demandes de financements peut rester inchangé voir augmenter malgré des simplifications en apparence drastiques : la division par deux de la taille des dossiers en Australie ayant augmenté la charge de travail lié aux demandes de financement (Barnett, Graves, Clarke, & Herbert, 2015). Une autre option serait la répartition égalitaire des financements. Pour donner un simple exemple, dissoudre l'Agence nationale de la recherche (ANR) et répartir équitablement son budget offrirait plus de 10k€ par an par EC du secteur public français. Ce montant est déjà plusieurs fois supérieur aux financements réguliers de la plupart des EC, équivalent par exemple à un financement de Projet exploratoire premier soutien (PEPS) pour chaque EC. Naturellement, ces financements pourraient être répartis différemment selon les domaines et les besoins — les dépenses de fonctionnement en biologie étant plus élevées qu'en mathématiques ou en droit. L'ANR n'est pas la seule source de financements compétitifs, et une étude sur d'autres pays européens montre qu'une répartition égalitaire du budget total mènerait à des montants beaucoup plus élevés (~32k€ par an aux Pays-Bas et un peu plus au Royaume-Uni, en sus du recrutement de personnes en doctorat ou post-doctorat au même rythme qu'aujourd'hui) (Vaesen & Katzav, 2017). Cette suggestion serait draconienne mais — en gardant comme objectif une augmentation de la productivité — une répartition plus égalitaire des crédits de recherche est soutenue par de nombreuses études empiriques (Aagaard, Kladakis, & Nielsen, 2020).

D'autres options méritent un examen plus approfondi, tels que des systèmes de loterie où une proportion constante des projets sont financés — potentiellement après une évaluation minimale visant à éliminer les projets irréalistes (Roumbanis, 2019). Cela peut sembler injuste pour les EC les plus performants mais, comme vu plus haut, de nombreux systèmes d'évaluation ne fonctionnent aujourd'hui que marginalement mieux que des loteries (Cortes & Lawrence, 2021). Ces dernières permettent par ailleurs d'éviter certains écueils comme la concentration des financements sur les projets les plus conservateurs (Avin, 2015). Un système de ce type a récemment été mis en oeuvre en Nouvelle-Zélande, le principe et les effets ayant été majoritairement acceptés par les EC qui en dépendent — qu'ils aient obtenu un financement ou non (Liu et al., 2020). Enfin, d'autres alternatives plus innovantes et potentiellement plus efficaces ont aussi été imaginées. Une proposition audacieuse serait d'attribuer à chaque EC un budget régulier dont la moitié doit être redistribuée anonymement vers ses pairs selon son jugement. Cela créerait une forme d'évaluation par les pairs fluide et distribuée, où le pouvoir limité conféré à chaque membre offre de meilleures garanties que le système actuel face à la possibilité

26. Le budget de fonctionnement de l'Agence Nationale de la Recherche est de l'ordre de 5% des crédits alloués, mais cela ne compte que les ~323 agents et pas les évaluateurs extérieurs.

de cartels et d'ententes entre EC (Bollen, Crandall, Junk, Ding, & Börner, 2017).

3 Conclusion : le productivisme est-il un dogmatisme ?

Nous avons vu qu'aucun des trois postulats annoncés en début d'article ne reposait à la fois sur une solide base scientifique et une confirmation empirique : les évaluations d'articles et de projets ont une part de hasard qui est parfois totale ; il n'est nullement certain que les mécanismes sélectifs fassent émerger les « meilleurs » ; l'attribution de moyens supplémentaires à ces derniers est inefficace voire contre-productive. Aucun indicateur fiable ne montre que les politiques élitistes et productivistes augmentent réellement la productivité, alors que leurs coûts se font de plus en plus sentir. Ces critiques ne sont pas neuves et s'accroissent depuis plusieurs décennies alors même que les politiques productivistes se normalisent, poussées par la rhétorique de l'urgence. Il semble donc approprié de chercher des explications en dehors des considérations purement empiriques.

L'analyse des systèmes de production et des institutions et de leurs discours révèle ainsi des divergences entre les objectifs affichés et les moyens mis en oeuvre pour les réaliser. Ces divergences peuvent d'abord s'expliquer par le détournement ponctuel des instances dirigeantes aux profit d'intérêts particuliers. Mais elles proviennent surtout de contradictions dans l'organisation du système, lesquelles reposent sur des présupposés d'ordre idéologiques enracinés dans l'histoire de la modernité socio-technique et désormais confondues avec le « bon sens » (Muller, 2017).

Plus précisément, le productivisme est de longue date un dogme de la modernité qui a transcendé les clivages politiques au XIX^e et XX^e siècles (Audier, 2019). La recherche sur la productivité naît d'abord dans l'industrie, dans une démarche se considérant comme apolitique et dont l'organisation scientifique du travail et le fordisme furent les résultats les plus visibles, dans une entreprise globale de libération des forces productives. La société industrielle de cette époque pouvait mesurer sa productivité en poids, volume et valeur rapportés aux moyens, sans que le productivisme américain ne soit fondamentalement différent du productivisme soviétique. Or, le *complexe technoscientifique* (Pickstone, 2001) de l'ère industrielle a laissé place à celui de l'ère de l'information et de la communication, marqué par la tertiarisation de l'économie et la complexification de la plupart des tâches productives. Le besoin de mesurer la productivité n'a pas disparu mais de plus en plus d'emplois résistent aux mesures simples qui guidaient l'optimisation des chaînes de montage. On voit donc le développement d'outils managériaux hautement complexes et de politiques publiques de plus en plus co-construites avec des acteurs privés (Rival & Ruano-Borbalan, 2017). La difficulté de mesure et le flou croissant entourant la production des emplois tertiaires réduisent aussi la possibilité d'observer directement et dans des temporalités courtes les conséquences des nouvelles organisations du travail, qu'elles soient guidées par des politiques publiques ou par des considérations managériales à l'échelle d'une PME. Paradoxalement, le renforcement constant de l'aiguillon compétitif crée une urgence de mesurer des productions parfois floues. Ce flou s'associe à un oubli — ou à une disparition — de certaines leçons du passé. Un exemple notable concerne la question du temps de travail. Malgré la durée légale française (35h hebdomadaires), les dépassements sont courants, y compris dans des secteurs à la fois techniques et critiques où l'erreur est coûteuse, comme l'hôpital public²⁷. Or les arguments principaux pour la réduction du temps de travail portent généralement sur trois facteurs :

- le besoin d'équilibre de vie et le confort des employés ;
- la démesure de la production dans le contexte de crise écologique ;
- la répartition du travail entre plus d'employés afin de réduire l'impact du chômage.

À ces arguments sont opposés des argumentations fondamentalement moralistes où le travail figure comme ordalie méritoire et où la crise économique requiert des efforts supplémentaires de la part des travailleurs. Les aspects idéologiques sont prépondérants des deux côtés du débat, laissant de côté la considération suivante : un employé travaillant plus d'heures produit-t-il plus ? Les travaux à ce sujet existent depuis l'époque de Ford (Crowden, 1930) et montrent que le nombre d'heures permettant une production maximale n'est pas forcément au dessus de 30-40h de travail hebdomadaire, dans des domaines multiples (Okazaki et al., 2019 ; Pang, 2016).

27. <https://www.legifrance.gouv.fr/ceta/id/CETATEXT000045959613>

L'économie de la connaissance n'est cependant pas nécessairement soumise aux mêmes règles. Dans le nouveau paradigme reposant sur l'innovation permanente, nombre d'acteurs de la recherche se calquent sur le fonctionnement des hauts centres de l'innovation. Ces lieux, tels que la Silicon Valley, cultivent une image du PDG génie dont la force de travail lui a permis de se construire en partant de rien. Ces figures exceptionnelles²⁸ sont des modèles censés être émulés par leurs employés, dans une perspective méritocratique et individualiste où ces derniers sont responsables de leurs situations — amplifiée par et amplifiant le manque de syndicalisation. On y voit donc des temps de travail hebdomadaires extrêmement élevés (Pfeffer, 2018), facilités par des politiques d'entreprises visant à maximiser la présence au travail en offrant de nombreux services à leurs employés (repas gratuits, divertissements et salles de sport sur site...). Il semblerait étrange que ces entreprises fortement dotées en services de ressources humaines ignorent l'impact du temps de travail sur la productivité (Pfeffer, 2001) ; d'autres logiques sous-jacentes sont donc probables. Celles-ci pourraient avoir trait d'une part au rapport social entre direction et employés, avec une vision de jeu à somme nulle : ce qui est bon pour l'employé est forcément mauvais pour l'entreprise, cette vision étant même visible chez certains défenseurs de la réduction du temps de travail²⁹.

L'émulation de ces pratiques temporelles — par exemple par le président de la République Emmanuel Macron³⁰ — n'épargne pas le secteur de la recherche. Ce dernier est probablement d'autant plus influencé par ce modèle que le système étasunien favorise les interactions entre recherche publique et grandes entreprises technologiques, symboles de l'excellence et de l'applicabilité des produits de la recherche.

Pour revenir à notre question centrale, nous avons montré que l'application à la recherche d'une logique productiviste résultant en une obsession quantitativiste reposait sur des postulats contredits par des données empiriques. Le « sens commun » très présent dans les instances de direction invite à évaluer continuellement les chercheurs au prix d'une part croissante de leur temps afin de savoir auxquels attribuer de plus ou moins importantes ressources. Les données empiriques montrent au contraire que cette logique est stérile voire nuisible, au moins dans le domaine de la recherche. Nous invitons à considérer les fondements idéologiques et culturels de la persistance des institutions dans ce fonctionnement, qui montrent que la naïveté n'est pas toujours où on l'attend. Il est temps de redonner à la science sa part dans le débat public sur la productivité, sans se fier au « bon sens » dogmatique.

28. Tim Cook, PDG d'Apple, propose le réveil à 4h du matin comme moyen d'améliorer le bien-être et la productivité.

29. Dans son ouvrage sur le repos et le travail, Pang (2016) observe des entreprises dont la productivité totale augmente suite à des réductions parfois drastiques du temps de travail. Faces à des crises conjoncturelles, ces dernières reviennent cependant sur ces mesures — qui augmentaient leurs bénéfices. L'auteur concède que la crise requiert des sacrifices de la part de toutes les parties, sans observer que le sacrifice en question est contre-productif. Là où une baisse des salaires augmenterait les marges de l'entreprise mais serait difficile pour les employés, une augmentation du temps de travail à salaire constant est tolérable, même si celle-ci ne réduit pas les coûts et diminue la productivité totale.

30. https://www.lepoint.fr/politique/textos-rendez-vous-secrets-marches-les-nuits-d-emmanuel-macron-18-08-2022-2486697_20.php

Références

- Aagaard, K., Kladakis, A., & Nielsen, M. W. (2020). Concentration or dispersal of research funding? *Quantitative Science Studies*, 1(1), 117–149.
- Aghion, P., & Cohen, É. (2004). *Éducation et croissance*. La documentation française, Paris.
- American Society for Cell Biology. (2012). San Francisco declaration on research assessment (DORA).
- Anonymous. (2010). The ratings game. *Nature*, 464. Consulté sur <https://www.nature.com/articles/464007b>
- Arfaoui, M. (2019). «creative europe» ou l'action publique par projet. enquête sur une modernisation des politiques publiques en europe. (Thèse de doctorat non publiée). EHESS-Paris.
- Audier, S. (2019). *L'âge productiviste : hégémonie prométhéenne, brèches et alternatives écologiques*. La Découverte.
- Auranen, O., & Nieminen, M. (2010). University research funding and publication performance—an international comparison. *Research policy*, 39(6), 822–834.
- Australian National Health and Medical Research Council. (2021). *Investigator grants 2020 outcomes factsheet* (Rapport technique). NHMRC. Consulté sur https://www.nhmrc.gov.au/sites/default/files/documents/attachments/grantdocuments/investigator_grants_2020_outcomes_factsheet.pdf
- Avin, S. (2015). Funding science by lottery. In *Recent developments in the philosophy of science : Epsa13 helsinki* (pp. 111–126). Springer.
- Ayoubi, C., Pezzoni, M., & Visentin, F. (2019). The important thing is not to win, it is to take part : What if scientists benefit from participating in research grant competitions? *Research Policy*, 48(1), 84–97.
- Azoulay, P. (2012). Turn the scientific method on ourselves. *Nature*, 484(7392), 31–32.
- Banal-Estañol, A., Macho-Stadler, I., & Pérez-Castrillo, D. (2019). Evaluation in research funding agencies : Are structurally diverse teams biased against? *Research Policy*, 48(7), 1823–1840.
- Barats, C. (2019). Médiatisation et usages socio-discursifs du «classement de Shanghai». *Revue française d'administration publique*(1), 101–120.
- Barnett, A. G., Graves, N., Clarke, P., & Herbert, D. (2015). The impact of a streamlined funding application process on application time : two cross-sectional surveys of australian researchers. *BMJ open*, 5(1), e006912.
- Baxt, W. G., Waeckerle, J. F., Berlin, J. A., & Callahan, M. L. (1998). Who reviews the reviewers? feasibility of using a fictitious manuscript to evaluate peer reviewer performance. *Annals of emergency medicine*, 32(3), 310–317.
- Biagioli, M., & Lippman, A. (2020). *Gaming the metrics : Misconduct and manipulation in academic research*. Mit Press.
- Bol, T., de Vaan, M., & van de Rijt, A. (2018). The Matthew effect in science funding. *Proceedings of the National Academy of Sciences*, 115(19), 4887–4890.
- Bollen, J., Crandall, D., Junk, D., Ding, Y., & Börner, K. (2017). An efficient system to fund science : from proposal review to peer-to-peer distributions. *Scientometrics*, 110(1), 521–528.
- Bornmann, L., Mutz, R., & Daniel, H.-D. (2010). A reliability-generalization study of journal peer reviews : A multilevel meta-analysis of inter-rater reliability and its determinants. *PloS one*, 5(12), e14331.
- Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, 71(3), 425.
- Chamayou, G. (2009). Petits conseils aux enseignants-chercheurs qui voudront réussir leur évaluation. *Contretemps*(1), 208–226. Consulté sur <http://www.contretemps.eu/petits-conseils-enseignants-chercheurs-qui-voudront-reussir-leur-evaluation/>
- Chaudret, B. (2013). Où en est la recherche française? *La Pensée*(4), 57–68. Consulté sur <https://www.cairn.info/publications-de-Bruno-Chaudret--705122.htm>
- Chawla, D. S. (2018). What's wrong with the journal impact factor in 5 graphs. *Nature Index*. Consulté sur <https://www.nature.com/nature-index/news-blog/whats-wrong-with-the-jif-in-five-graphs>
- Chawla, D. S. (2021). Scientists at odds on Utrecht University reforms to hiring and promotion

- criteria. *Nature Index*. Consulté sur <https://www.nature.com/nature-index/news-blog/scientists-argue-over-use-of-impact-factors-for-evaluating-research>
- Cortes, C., & Lawrence, N. D. (2021). Inconsistency in conference peer review : revisiting the 2014 NeurIPS experiment. *arXiv preprint arXiv :2109.09774*.
- Crowden, G. (1930). Industrial efficiency and fatigue. *Proceedings of the Royal Society of Medicine*, *23*, 479–85.
- Demetrescu, C., Finocchi, I., Ribichini, A., & Schaerf, M. (2020). On bibliometrics in academic promotions : A case study in computer science and engineering in italy. *Scientometrics*, *124*(3), 2207–2228.
- Dericquebourg, R. (2001). Mystagogie et religions de guérison : Max weber revisité. *Archives de sciences sociales des religions*(113), 149–158.
- Desailly-Chanson, M.-A., Penaud, M., Vilboeuf, L., Elshoud, S., & Piozin, E. (2022). *Contrôle de l'ihu méditerranée infection, rapport définitif* (Rapport technique). Inspection générale des affaires sociales. Consulté sur https://www.igas.gouv.fr/IMG/pdf/rapport_ihu_definitif_tome_1_.pdf
- Dougherty, M. R., & Horne, Z. (2022). Citation counts and journal impact factors do not capture some indicators of research quality in the behavioural and brain sciences. *Royal Society Open Science*, *9*(8), 220334.
- Ebadi, A., & Schiffauerova, A. (2016). How to boost scientific production? A statistical analysis of research funding and other influencing factors. *Scientometrics*, *106*(3), 1093–1116.
- Fire, M., & Guestrin, C. (2019). Over-optimization of academic publishing metrics : observing goodhart's law in action. *GigaScience*, *8*(6), giz053.
- Franzoni, C., Scellato, G., & Stephan, P. (2011). Changing incentives to publish. *Science*, *333*(6043), 702–703.
- Garcia-Costa, D., Squazzoni, F., Mehmani, B., & Grimaldo, F. (2022). Measuring the developmental function of peer review : a multi-dimensional, cross-disciplinary analysis of peer review reports from 740 academic journals. *PeerJ*, *10*, e13539.
- Griesemer, J. (2020). Gaming the metrics. In (p. 77-87). UC Davis Books.
- Gross, K., & Bergstrom, C. T. (2019). Contest models highlight inherent inefficiencies of scientific funding competitions. *PLoS biology*, *17*(1), e3000065.
- Hennessey, B. A., & Amabile, T. M. (2010). Creativity. *Annual review of psychology*, *61*, 569–598.
- Herbert, D. L., Barnett, A. G., Clarke, P., & Graves, N. (2013). On the time spent preparing grant proposals : an observational study of Australian researchers. *BMJ open*, *3*(5), e002800.
- Herbert, D. L., Coveney, J., Clarke, P., Graves, N., & Barnett, A. G. (2014). The impact of funding deadlines on personal workloads, stress and family relationships : a qualitative study of Australian researchers. *BMJ open*, *4*(3), e004462.
- Herbert, D. L., Graves, N., Clarke, P., & Barnett, A. G. (2015). Using simplified peer review processes to fund research : a prospective study. *BMJ open*, *5*(7), e008380.
- Jacob, B. A., & Lefgren, L. (2011). The impact of research grant funding on scientific productivity. *Journal of public economics*, *95*(9-10), 1168–1177.
- Larivière, V., & Gingras, Y. (2010). The impact factor's Matthew effect : A natural experiment in bibliometrics. *Journal of the American Society for Information Science and Technology*, *61*(2), 424–427.
- Larregue, J., & Saint-Martin, A. (2019). Troubles dans l'ethos scientifique. *Zilsel*, *6*(2), 149–163.
- Laudel, G. (2006). The art of getting funded : how scientists adapt to their funding conditions. *Science and Public Policy*, *33*(7), 489–504.
- Lauer, M. S., Danthi, N. S., Kaltman, J., & Wu, C. (2015). Predicting productivity returns on investment : thirty years of peer review, grant funding, and publication of highly cited papers at the national heart, lung, and blood institute. *Circulation research*, *117*(3), 239–243.
- Liu, M., Choy, V., Clarke, P., Barnett, A., Blakely, T., & Pomeroy, L. (2020). The acceptability of using a lottery to allocate research funding : a survey of applicants. *Research integrity and peer review*, *5*(1), 1–7.
- Madsen, E. B., & Aagaard, K. (2020). Concentration of danish research funding on individual research

- chers and research topics : Patterns and potential drivers. *Quantitative Science Studies*, 1(3), 1159–1181.
- Maisonneuve, H., Plaud, B., & Caumes, E. (2020). Pandémie à SARS-CoV-2 : éthique et intégrité oubliées devant la précipitation pour publier. *La Presse Médicale Formation*, 1(6), 572–581.
- McKiernan, E. C., Schimanski, L. A., Nieves, C. M., Matthias, L., Niles, M. T., & Alperin, J. P. (2019). Meta-research : Use of the journal impact factor in academic review, promotion, and tenure evaluations. *Elife*, 8, e47338.
- Mondello, G. (2022). Construction de systèmes de croyances et controverses autour du Covid-19. *Revue d'économie politique*, 132(1), 79–111.
- Mongeon, P., Brodeur, C., Beaudry, C., & Larivière, V. (2016). Concentration of research funding leads to decreasing marginal returns. *Research Evaluation*, 25(4), 396–404.
- Mullard, A. (2021). Half of top cancer studies fail high-profile reproducibility effort. *Nature*, 600(7889), 368–369.
- Muller, S. M. (2017). Academics as rent seekers : Distorted incentives in higher education, with reference to the south african case. *International Journal of Educational Development*, 52, 58–67.
- National Science Board. (2014). *Reducing investigators' administrative workload for federally funded research* (Rapport technique). National Science Foundation. Consulté sur <https://www.nsf.gov/pubs/2014/nsb1418/nsb1418.pdf>
- Obadia, L. (2021). Urgence de la recherche, recherche en urgence. *Questions de communication*, 40(2), 213–232.
- Okazaki, E., Nishi, D., Susukida, R., Inoue, A., Shimazu, A., & Tsutsumi, A. (2019). Association between working hours, work engagement, and work productivity in employees : A cross-sectional study of the japanese study of health, occupation, and psychosocial factors relates equity. *Journal of Occupational Health*, 61(2), 182–188.
- Pang, A. S.-K. (2016). *Rest : Why you get more done when you work less*. Basic Books.
- Park, S., & Berry, F. (2014). Successful diffusion of a failed policy : The case of pay-for-performance in the us federal government. *Public Management Review*, 16(6), 763–781.
- Patience, G. S., Patience, C. A., Blais, B., & Bertrand, F. (2017). Citation analysis of scientific categories. *Heliyon*, 3(5), e00300.
- Pfeffer, J. (2001). What's wrong with management practices in Silicon Valley? A lot. To prevent high turnover, burnout and loss of employee commitment, learn to avoid four practices that are undermining some high-profile companies. *MIT Sloan Management Review*, 42(3), 101–103.
- Pfeffer, J. (2018). Work hours and health : A comment on “beyond nine to five”. *Academy of Management Discoveries*, 4(1), 94–96.
- Pickstone, J. V. (2001). *Ways of knowing : A new history of science, technology, and medicine*. University of Chicago Press.
- Pier, E. L., Brauer, M., Filut, A., Kaatz, A., Raclaw, J., Nathan, M. J., ... Carnes, M. (2018). Low agreement among reviewers evaluating the same nih grant applications. *Proceedings of the National Academy of Sciences*, 115(12), 2952–2957.
- Piller, C. (2022). Blots on a field? *Science*, 377(6604). Consulté sur <https://www.science.org/content/article/potential-fabrication-research-images-threatens-key-theory-alzheimers-disease>
- Quan, W., Chen, B., & Shu, F. (2017). Publish or impoverish : An investigation of the monetary reward system of science in China (1999-2016). *Aslib Journal of Information Management*.
- Reymert, I. (2021). Bibliometrics in academic recruitment : A screening tool rather than a game changer. *Minerva*, 59(1), 53–78.
- Reymert, I., Jungblut, J., & Borlaug, S. B. (2021). Are evaluative cultures national or global? a cross-national study on evaluative cultures in academic recruitment processes in europe. *Higher Education*, 82(5), 823–843.
- Rival, M., & Ruano-Borbalan, J.-C. (2017). La fabrique des politiques publiques co-construites : idéologie et pratiques innovantes. *Politiques et management public*, 34(1-2), 5–16.
- Rodamar, J. (2018). There ought to be a law! campbell versus goodhart. *Significance*, 15(6), 9-9.

- Consulté sur <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1740-9713.2018.01205.x> doi: <https://doi.org/10.1111/j.1740-9713.2018.01205.x>
- Rodgers, P., & Collings, A. (2021). Reproducibility in cancer biology : What have we learned? *Elife*, *10*, e75830.
- Roebber, P. J., & Schultz, D. M. (2011). Peer review, program officers and science funding. *PLoS One*, *6*(4), e18680.
- Roumbanis, L. (2019). Peer review or lottery? a critical analysis of two different forms of decision-making mechanisms for allocation of research grants. *Science, Technology, & Human Values*, *44*(6), 994–1019.
- Sîle, L., & Vanderstraeten, R. (2019). Measuring changes in publication patterns in a context of performance-based research funding systems : the case of educational research in the University of Gothenburg (2005–2014). *Scientometrics*, *118*(1), 71–91.
- Stephan, P., Veugelers, R., & Wang, J. (2017). Reviewers are blinkered by bibliometrics. *Nature*, *544*(7651), 411–412.
- Stewart, M. (2018). State crime by proxy : Corporate influence on state sanctioned social harm. *Journal of Critical Psychology, Counselling and Psychotherapy Vol, 18*, 217–227.
- Taffe, M. A., & Gilpin, N. W. (2021). Equity, diversity and inclusion : Racial inequity in grant funding from the us national institutes of health. *Elife*, *10*, e65697.
- Teixeira da Silva, J. A., & Dobránszki, J. (2017). Highly cited retracted papers. *Scientometrics*, *110*(3), 1653–1661.
- Vaesen, K., & Katzav, J. (2017). How much would each researcher receive if competitive government research funding were distributed equally among researchers? *PLoS One*, *12*(9), e0183967.
- Van Dijk, D., Manor, O., & Carey, L. B. (2014). Publication metrics and success on the academic job market. *Current Biology*, *24*(11), R516–R517.
- Wahls, W. P. (2018a). High cost of bias : Diminishing marginal returns on NIH grant funding to institutions. *BioRxiv*, 367847.
- Wahls, W. P. (2018b). Point of view : The NIH must reduce disparities in funding to maximize its return on investments from taxpayers. *Elife*, *7*, e34965.