

Mesurer la performance des chercheurs, au risque de la bureaucratie

Nicolas Berland
Professeur d'Université
Université Paris-Dauphine, DRM

Benjamin Drevetton
Maitre de conférences
IAE de Poitiers, CEREGE

Introduction

Nous sommes entrés dans l'ère des sociétés de la connaissance qui nous amène à considérer les activités de recherche avec une acuité nouvelle. Exercée jusqu'alors en marge de la société productive, la recherche s'est vue demandée des comptes sur son efficacité et son utilité. Les chercheurs sont devenus « *accountable* » (Osterloh et al., 2008). Des pratiques de contrôle qui existaient déjà se sont durcies, professionnalisées et institutionnalisées. Dans le cadre de cet article, nous chercherons à convaincre le lecteur qu'elles se sont bureaucratisées, c'est-à-dire que les règles de contrôle l'ont parfois emporté sur les finalités du contrôle. La logique de fonctionnement l'emporte sur la logique de mission et déclenche des effets secondaires pervers. En décortiquant les mécanismes pervers de ce contrôle nous montrerons qu'apparaissent les critiques usuelles des processus de mise sous tension de la performance. Nous concluons que, malgré tout, ces dérives ne sont pas inévitables.

Les logiques de contrôle qui se sont développées font appel à une large gamme des outils de mesure de la performance : indicateurs d'activité et de performance, budgets organisés par finalités, comptabilité de gestion... et s'inscrivent donc dans le cadre du développement du *New Public Management* (NPM) ou nouvelle gestion publique (Hood 1991 et 1995) lancé depuis quelques années dans le monde anglo-saxon et qui a trouvé une résonance particulière en France avec le vote de la Lof (Loi organique relative aux Lois de Finance) en 2001 et plus récemment la RGPP (Revue Générale des Politiques Publiques). L'université n'échappe pas à ce mouvement de managérialisation. La loi n° 2007-1199 du 10 août 2007 relative aux Libertés et Responsabilités des Universités (LRU) prévoit que, d'ici 2013, toutes les universités accèdent à l'autonomie dans les domaines budgétaire (article 50) et de gestion des ressources humaines. Cette nouvelle maîtrise de la masse salariale engendre l'implantation de nouveaux outils de gestion afin que gestionnaires et décideurs politiques puissent assurer un pilotage opérationnel et stratégique de leur UFR (Chatelain-Ponroy et al. 2006). Aussi, ces processus d'instrumentation constituent de véritables ruptures pour des organisations traditionnellement régulées par des normes juridiques.

Dans cet article, nous examinerons les outils de contrôle et de comptabilité mis en place pour juger de l'activité de recherche avec un regard critique mais bienveillant. Critique car nous pensons qu'ils font parfois l'objet d'une application technocratique par des acteurs ayant tendance à les considérer trop rapidement comme de simples artefacts techniques dénués de valeur et d'ancrage social. Or, les outils comptables sont tout sauf neutres et incorporent des valeurs, des croyances et des idéologies qu'ils empruntent à leur environnement institutionnel. De ce fait, décortiquer les outils de comptabilité revient à questionner leurs fondements sociaux. Les qualités d'objectivité des outils de contrôle sont alors mises à mal et ces outils ne reflètent sans doute que des vues partisans, partiales et partielles. Pour autant, notre critique est bienveillante dans la mesure où nous pensons qu'il est légitime de contrôler et de mesurer la performance des activités de recherche. En effet, la massification de l'enseignement supérieur rend nécessaire un contrôle à distance. La recherche n'est plus une activité marginale, cantonnée à quelques individus. Aussi, nous tenterons de faire des propositions pour une évaluation apaisée de la performance de la recherche. Nous nous demanderons alors comment évaluer les activités de recherche académiques ? Nous devons toutefois d'emblée circonscrire notre propos. Nous ne traiterons que de la recherche en sciences de gestion, dans un contexte français. En effet, ce papier est un travail réflexif lié à notre vécu de chercheurs en sciences de gestion étudiant les pratiques du pilotage de la performance des organisations. Certes, nous connaissons la situation des sciences de gestion à l'étranger (en fait le monde

anglo-saxon !) ou dans d'autres champs académiques mais ces différentes situations n'ont pas pour nous la même familiarité. Dans quelle mesure ces réflexions et propositions sont-elles transposables ailleurs, nous laissons le lecteur en juger par lui-même.

Dans une première partie, nous tenterons de brosser un état des lieux de l'évaluation comptable de la recherche pour ensuite nous interroger sur quelques unes des causes des évolutions constatées ces dernières années. La seconde partie sera consacrée à mettre en évidence les impasses, pas forcément inévitables, dans lesquelles nous conduisent certaines pratiques. Enfin, la dernière partie visera à faire des propositions pour une évaluation apaisée de la performance de la recherche en sciences de gestion.

1. L'état des lieux de l'évaluation de la recherche

Depuis 2006, le gouvernement a décidé de mesurer la production scientifique prenant acte de la faible publication de nombre d'enseignants-chercheurs. Les sciences de gestion n'échappent pas à ce phénomène. Aussi a-t-on vu la multiplication d'instruments de mesure. Après avoir présenté les principaux éléments caractéristiques de cette dynamique (1.1), nous nous intéresserons aux raisons de cette évolution (1.2).

1.1. Une volonté d'objectiver la recherche

Les enseignants chercheurs ont une triple activité d'enseignement, de recherche et de gestion administrative (des diplômes, des centres de recherche...). L'activité de recherche est une des singularités de l'enseignement dans le supérieur (Université ou Ecole de commerce). C'est parce que l'enseignant-chercheur a une activité réflexive sur la discipline qu'il enseigne (étude des pratiques dans le cadre de protocoles scientifique permettant de distinguer les effets de mode des tendances durables par exemple) qu'il acquiert une légitimité à enseigner à haut niveau. Sans cette dernière, il est le répétiteur de savoirs conçus ailleurs. La recherche lui confère une légitimité et une originalité qui le distingue de l'enseignant. Il est d'usage de considérer que cette activité l'occupe en général la moitié de son temps. Après avoir identifié les instances de contrôle de cette activité, nous reviendrons sur les caractéristiques de l'activité de recherche et sur les outils permettant d'objectiver ses résultats.

Pour contrôler cette activité, de nombreuses procédures fonctionnent depuis des années. A l'Université, le CNU (Conseil National des Universités) ou l'Agrégation du supérieur sanctionnent les bonnes performances dans le cadre de mécanismes de promotion vers les plus hautes fonctions du professorat. Dans les Ecoles, des mécanismes d'évaluation régulière (par ex. le mécanisme des *tenure tracks*) jouent le même rôle. Plus récemment, l'AERES (Agence de l'Evaluation de la Recherche et de l'Enseignement Supérieur) a été créé afin de mettre en place une procédure d'évaluation des unités de recherche. Indirectement, il s'agit bien là aussi de contrôler l'activité du chercheur. Pour autant ces procédures ne sont pas exemptes de défauts dès lors que le jeu politique des acteurs devient prépondérant et assure la promotion de mauvais candidats ou quand elles ne s'intéressent qu'à une minorité d'individus, ceux ayant choisi d'essayer d'obtenir une promotion en se jouant du système bureaucratique de l'évaluation.

Aussi a-t-on vu se développer depuis quelques années de nouveaux instruments de mesure associés à un nouveau vocabulaire. Comment qualifier un « bon chercheur » ? Pour répondre à cette interrogation, plusieurs qualificatifs ont été utilisés pour désigner l'activité de

recherche. Ce sont succédés dans le temps, le « chercheur actif », le chercheur « produisant », le chercheur « publiant ». Aujourd'hui, il semble que le bon chercheur soit l'enseignant-chercheur « produisant en recherche et valorisation ». Ce chercheur est tout d'abord un chercheur publiant : l'enseignant-chercheur « produisant en recherche et valorisation » est celui qui « *dans le cadre d'un contrat quadriennal, satisfait à un nombre minimal de « publications »*. Il doit sur une période donnée (jusqu'ici quatre ans) publier dans des revues scientifiques, c'est-à-dire des revues ayant un processus d'acceptation des articles se faisant sur la base d'une révision en double aveugle. Le bon chercheur affiche alors un score de publication. Mieux encore, plus qu'écrire, le bon chercheur est lu (mais cela est difficile à mesurer) et surtout il est cité dans d'autres revues. Ainsi, au « chercheur publiant » s'additionne le « chercheur reconnu ». Ce dernier doit montrer qu'au-delà de ses publications, il s'insère dans sa communauté scientifique et participe activement à son dynamisme. Aussi, afin de juger de cette implication, un ensemble d'outils permet alors de mesurer la notoriété du chercheur soit dans des bases comme le SSCI (*Social Science Citation Index*), soit grâce à des logiciels comme Harzing (*Publish or Perish*) qui répertorie sur Google Scholar les documents faisant référence aux articles du chercheur. Le bon chercheur se voit alors affecter des scores de publication. Au premier temps des classements, il accumulait des étoiles pour se rendre *bankable* (sic) tel un joueur de foot participant au *mercato* des profs (voire *l'Expansion*, mai 2011). De façon plus sophistiquée, plus moderne mais aussi plus obscure, il peut être pourvu d'un h-index et d'un g-index qui synthétisent sa performance en notoriété. En synthèse, le bon chercheur est donc publiant et s'engage dans des tâches d'intérêt collectif qui dynamise sa communauté de recherche de référence.

1.2. Les raisons de cette évolution

Pourquoi en est-on arrivé là ? Les raisons sont multiples et nous en retiendront quatre.

Tout d'abord, il nous semble que la massification de longue durée de l'enseignement supérieur (un mouvement commencé avant 1968) nécessite de revoir les modes d'évaluation. Quand l'activité de recherche est limitée à un petit nombre d'individus, non seulement l'évaluation individu par individu est possible sur la base du contenu de leurs travaux, mais l'absence d'évaluation est aussi une option car le coût du contrôle dépasserait alors les bénéfices que l'on peut en attendre. La situation est très différente avec une population de chercheurs ayant fortement augmenté. Non seulement, le traitement individuel n'est plus possible et un traitement plus automatique, à distance et moins coûteux s'impose. Néanmoins, le coût d'une population de chercheurs inefficace est très élevé pour la population. On peut même se demander si la qualification de chercheurs s'applique et doit s'appliquer à des intervenants du supérieur qui font face à une demande sociale allant du BTS et l'IUT au doctorat. Le centre de gravité des populations ayant besoin d'enseignant-chercheur ne s'est-il pas déplacé sans que le statut des personnes concernées n'ait véritablement évolué ? Dans ce cas, la pression mise à la publication vise, à moyen terme, à recadrer un dispositif sans pour autant s'attaquer au statut, via par exemple la création de collèges universitaires chargés d'enseignement dans certains cas, entourés de quelques pôles de recherche dans d'autres.

La deuxième raison est le souci de contrôler les connaissances diffusées afin de mieux former les cadres de demain. Sans un contrôle strict des connaissances produites et valides, le risque est fort de voir se développer des enseignements de complaisance, fondés sur des modes managériales. Mais comment contrôler les connaissances produites ? La soumission de ces recherches dans des revues est sans doute l'un des moyens les plus efficaces (à l'exclusion des

autres ?) de contrôler la qualité produite. Encore une fois, c'est la massification du marché de l'enseignant-chercheur qui nécessite de s'assurer du contrôle des connaissances. Dans ce cadre, la publication dans des revues sélectives peut être soumise à une autre interprétation. Il ne s'agit plus de production de connaissances mais de garantir que le corps enseignant se maintient à un niveau de vigilance sur les pratiques, qui soit compatible avec les objectifs de qualité.

Troisièmement, cette évolution a été rendue possible par l'institutionnalisation du champ académique. Pour compter des publications, il faut des revues, des comités de lecture, des habitudes d'évaluation qui ne sont pas si anciennes en France. Avant les années 1990, le *Revue Française de Gestion* constituait l'un des rares supports de publication par exemple. En comptabilité, la création de la revue *Comptabilité Contrôle Audit* remonte à 1995, celle de la revue *Finance Contrôle et Stratégie* à 1998.

Enfin, il faut noter que l'idée de mesurer la production du chercheur s'inscrit dans un contexte social propice à une idéologie de la mesure qui n'a cessé de monter depuis trente ans, en s'accélégrant depuis peu. « *On ne gère bien que ce que l'on mesure* » dit l'adage. Si cette proposition est très critiquable (Berland et al. 2008), elle a pourtant été reprise dans des champs qui dépassent largement le cadre de l'entreprise. Notre société est celle du *ranking* (voir, par exemple, les classements effectués sur les diplômés) et des objectifs qui ne sauraient mentir sur la performance de l'évalué. Les chercheurs ne sont pas les seuls acteurs publics à être soumis à une évaluation de leurs activités. Ainsi, en janvier 2008, le porte-parole du gouvernement annonçait la mise en place 30 « critères chiffrés de performance » pour évaluer les résultats de l'action des ministres¹. A l'instar de la trajectoire observée chez les médecins, verra-t-on à l'avenir l'apparition d'un nouveau type de chercheur chargé de répertorier l'activité des membres de son laboratoire de recherche ?

II – La naissance d'une nouvelle figure d'acteur : le chercheur auditable

Le chercheur devient évaluable. A la suite de ce constat apparaît la naissance d'un nouveau « type » de chercheur : le chercheur auditable. Néanmoins, l'activité de ce dernier s'avère plus facile à qualifier (2.1) qu'à contrôler (2.2).

2.1 Un chercheur auditable

La recherche est donc devenue une activité comme une autre et dans laquelle il est possible de produire une *bottom line* de l'activité sous forme d'un résultat numérique. A l'instar de l'activité industrielle et commerciale, le chercheur est un publiant calculable (Miller et O'Leary, 1987).

Pour autant cette qualité ne va pas de soi et comme l'a montré Power (1999), une intense activité en amont est nécessaire pour permettre de mesurer « objectivement ». Pour être mesurable ou auditable, un système doit être organisé en amont pour avoir ces

¹ A titre d'exemple, la ministre de la Culture, Christine Albanel, devait être jugée sur « l'évolution de la fréquentation des musées lorsqu'ils sont gratuits » ou bien sur « la part de marché des films français en France ».

caractéristiques. Dans le cas de la recherche en sciences de gestion, que faut-il compter comme étant une publication ? Pour résoudre ce problème, des listes de revues scientifiques acceptables ont été produites. Les revues ont été classées selon leurs qualités supposées en différents niveaux d'excellence, faisant la part belle aux publications anglo-saxonnes, supposées, sans doute à raison mais pas toujours (Hopwood 2008), de meilleures qualités. De tels classements existaient déjà au niveau international, notamment diffusés par des journaux tels que le Financial Times. Leur usage était de ce fait répandu dans certaines Ecoles de commerce, les plus prestigieuses et jouant dans l'arène de la compétition internationale. La première version du référentiel publiée par le CNRS s'est faite en douceur. Mais dès que le potentiel d'intrusion dans l'activité des chercheurs et de leur laboratoire a été identifié, les éditions suivantes ont été plus compliquées et ont donné lieu à d'âpres négociations avec les différentes parties prenantes dont les sociétés savantes.

Conscient de certains effets pervers, ces listes n'ont cessé d'être réaménagées par les parties prenantes. La première liste attribuait ainsi des étoiles. Les « moins bonnes revues » (mais qui avaient quand même la chance d'être classées) avaient une étoile, les meilleures en ayant quatre ou cinq. Certains, dont la presse généraliste, ce qui montre l'importance qu'a pu prendre cette activité de classement, se sont empressés d'ajouter les étoiles. L'objectif des chercheurs était alors défini en nombre d'étoiles à atteindre sur une période donnée. Tout cela étant fait sans se soucier du contenu des articles en question et sans se préoccuper non plus de savoir s'il existait ou non une relation cardinale en lieu et place de la relation ordinale qui avait été imaginée au départ. Il est toutefois intéressant de constater que bien que le système d'étoiles ait été supprimé et le classement inversé pour éviter les additions qui n'ont pas de sens, le système continue à fonctionner dans beaucoup d'Ecoles de commerce (en général celles assez éloignées des préoccupations de la recherche) et dans les journaux. Il faut dire que le système, à défaut d'être juste et pertinent, est simple à comprendre et à commenter.

Et, c'est à l'aune de ces listes de revues que la productivité des chercheurs a été définie car il existait une norme permettant de comparer les différents laboratoires et les résultats de l'activité de ses individus. La vraie nouveauté tient plus à la liste qu'à la procédure de calcul elle-même. Sans ces listes, les procédures de calcul sont impossibles.

2.2 Les impensés de l'évaluation de la recherche en sciences de gestion

Tout ce travail de normalisation et de mesure laisse toutefois dans l'ombre des pans entiers de l'activité du chercheur. Passons sur le fait que le chercheur est souvent un enseignant-chercheur et qu'à ce titre, il exerce au moins deux autres missions, une d'enseignement et l'autre d'administration. La performance pour un individu donné ou pour une institution serait à rechercher dans l'équilibre entre ces trois activités et non dans une évaluation qui découplerait les activités les unes des autres. Mais l'équilibre reste un concept flou difficile à mesurer. Ainsi, la focalisation sur une de facette du métier d'enseignant chercheur aboutit à appauvrir la performance de ces individus. De plus, elle génère des attitudes individualistes qui ne vont pas à l'encontre de la vie de l'institution mais qui la limitent dans son champ d'action. Toutefois, nous pourrions croire que cloisonner les trois missions, et donc séparer l'activité de recherche, aboutit à une meilleure évaluation du chercheur. Or, comme nous allons tenter de l'illustrer, les systèmes de contrôle de l'activité de recherche aujourd'hui mis en place laissent de côté des pans entiers du métier de chercheur ...

Premier impensé : les spécificités des méthodes de recherche en sciences de gestion. L'objet de recherche des chercheurs en gestion étant l'organisation, qu'en est-il des activités de recherche appliquée que pratiquent certains chercheurs ? Si dans certaines disciplines, les

brevets interviennent dans la mesure de la performance, pourquoi les missions de recherche-action, de recherche-intervention ne compteraient-elles pas dès lors qu'elles sont exercées dans le cadre du laboratoire ? Les procédés de gestion sont rarement brevetables et les sciences de gestion se définissant comme des sciences de l'action (Argyris, 1995), il est regrettable de ne pas compter ces activités, fussent-elles différentes des activités de publication. L'enjeu est de taille quand ces activités rapportent 2 millions d'euros par an à un laboratoire alors que la dotation du ministère est de... 30 000 euros ! Si la recherche permet de s'assurer de la pertinence du contenu des matières enseignées, la recherche appliquée devrait permettre de son côté de s'assurer du contenu des connaissances pour lesquelles les organisations consultent.

Deuxième impensé de l'activité du chercheur, les supports de diffusion. Nous pensons ici principalement d'une part aux prises de parole dans les grands médias pour commenter l'actualité et, d'autre part à l'écriture d'ouvrages. La caricature est d'opposer journalisme et recherche. Mais avoir des chercheurs coupés à ce point des réalités qu'ils ne seraient pas capable d'intervenir dans les media serait tout aussi caricatural. Bien sûr, des interventions dans les media, des interventions en organisation ou des publications pédagogiques de chercheurs ne publiant jamais seraient suspectes. Un équilibre, une éthique de la recherche apparaissent plus complexes à évaluer que la réduction de l'activité à la simple mesure du nombre de publications. De la même façon, comment qualifier l'activité d'écriture d'ouvrages pédagogiques ? Un chercheur engagé dans un tel processus confirmera aisément que ce travail est fortement chronophage. Pourtant, il est surprenant de constater que les ouvrages ont longtemps été les grands absents des classements. Un récent rapport de la FNEGE semble toutefois ouvrir des perspectives (Commission pour l'évaluation et la valorisation du livre). Comment compter les livres ? Avant de répondre à cette question, il est amusant de noter que la question se pose sur les livres mais pas sur la production électronique, sites internet, podcasts..., étrange façon d'entrer dans l'histoire et de vivre avec son temps ! Mais gardons les livres tant leur potentiel académique est fort. Il est évident que la publication d'ouvrages concourt grandement à la diffusion de connaissances. Il est intéressant de noter sous Harzing que la notoriété vient plus souvent des livres que des articles. Bien sûr l'objection est que tous les livres ne se valent pas, mais nous tâcherons de répondre à cette objection plus tard. Ne pas retenir une production numérique au prétexte qu'on ne sait pas la compter pose des questions. Par exemple, cela amène à se demander pourquoi telle ou telle revue est prise en compte, comment s'effectue sont travail de sélection... Avec un peu de pratique de cet environnement, on découvre des choses au moins aussi étonnantes (développement des numéros spéciaux de revues, activation des réseaux sociaux etc.).

Enfin, troisième impensé : les cibles de la diffusion. Ainsi, qu'en est-il des activités de diffusion grand public de la connaissance ? Produire de la connaissance académique est nécessaire, la diffuser est impératif. Nombre de chercheurs produisent des publications pour des revues professionnelles ou à des fins pédagogiques. Ne pas répondre aux revues pédagogiques revient à enfermer les chercheurs dans leur tour d'ivoire, à l'heure où leurs idées peuvent recevoir une forme de validation externe, dans un contexte où elles sont souvent appréciées car rompant avec un conformisme très souvent taylorien. De même, la production d'études de cas à des fins pédagogiques est une nécessité dans une discipline qui se caractérise par de forts bataillons d'étudiants. Ne pas assumer ce rôle revient à toujours compter sur les mêmes mise en situation de nos étudiants devant des problèmes sans cesse rabâchés car facile à produire et laissant donc du temps pour accomplir ce qui est vraiment important, à savoir des études que personne ne lit vraiment.

Un constat s'impose : se focaliser sur la seule activité de recherche (évaluée/publiée) suffit à laisser dans l'ombre des activités qui pourraient pourtant, dans un registre élargi, être qualifiées de recherche. Ainsi, en se concentrant principalement sur l'activité d'un certain type de publications, l'évaluation de la performance des chercheurs se révèle étriquée, simplifiée.

Au final, l'objectivité de la mesure donne un caractère scientifique à des données et des contextes difficiles à saisir. Toutefois cette quantification de la recherche ne va pas sans poser de problème car elle nous permet d'oublier qu'une bonne mesure repose d'abord un ensemble de conventions solidement établies (Desrosières, 1993) et qui elles n'ont rien d'objectifs.

3. Mesurer la performance du chercheur, une activité sociale

Mettre en équation l'activité scientifique du chercheur se fait en s'intéressant aux *outputs* de son activité et non à la création de valeur qu'il produit, c'est-à-dire les *outcomes*. De ce fait, la mesure de la performance du chercheur conduit généralement à une impasse car elle ignore les conventions sociales de cette activité (3.1). Cette ignorance engendre mécaniquement des effets pervers (3.2). Pour autant, au-delà de cette dimension critique de la mesure de la performance du chercheur en sciences de gestion, ce sont précisément ces bases sociales qui permettent de reconstruire les outils d'une évaluation apaisée de la recherche (3.3).

3.1 Mise au jour de quelques conventions sous-jacentes à la mesure

L'analyse de l'activité peut permettre de mettre en évidence les trois conventions sur lesquelles repose la construction d'un indicateur à même d'évaluer les résultats d'une recherche.

La première convention est « sociale » : l'activité de recherche ne peut être analysée en dehors de son inscription sociale. Comme l'indique Argyris (1995), le chercheur ne peut être séparé du citoyen. Pourtant, l'activité du chercheur a longtemps été appréhendée comme un travail solitaire. Le chercheur lit, écrit et communique seul. Néanmoins, si nous analysons l'activité de publication, le constat inverse peut être établi. En effet, produire des articles de recherche n'est pas une activité solitaire se pratiquant seul face à son écran d'ordinateur. Publier est une activité sociale en réseau (Latour, 1979, 1987). Cette activité nécessite des alliés de diverses sortes. Pour être publié, il faut produire de bons articles. Mais cela ne suffit pas car il existe beaucoup de bons articles mais aussi parce qu'on ne sait pas toujours si l'article est bon ou pas. C'est souvent le fait d'être publié qui lui confère cette qualité. Il est donc préférable de bien maîtriser le jeu social de la recherche en étant capable d'aller dans les conférences qui comptent, celles où des éditeurs de revues vont vous remarquer (il existe d'autres façons de se faire remarquer d'eux : soutenance de thèse, invitation comme guest speaker...). Mais il faut aussi se faire repérer des potentiels réviseurs. En lisant votre papier (anonyme), ils seront ainsi réconfortés de savoir que ce dernier a « tourné » dans des colloques. Ils auront entendu d'autres personnes réagir sur votre papier, ce qui confortera leur jugement. Mais, le mieux est d'avoir un réseau si dense que beaucoup de vos alliés attendent votre article pour pouvoir le citer. Non par pure générosité mais parce que l'article cite leurs publications. Donc un bon publiant, bien cité a beaucoup d'alliés qui le rendent attractif sur le marché de l'emploi. Si en plus, ses papiers sont intéressants, alors il s'agit sans aucun doute

de bonne recherche. Mais la mesure reste celle de la force du réseau et non de la seule valeur scientifique.

La deuxième convention est « procédurale ». L'activité de recherche est, par essence, une activité de long terme. Or, il est intéressant de noter que l'évaluation de la recherche souffre d'une limite assez proche de l'évaluation financière d'une entreprise. Tout comme le ROI (*Return On Investment*) mesure mal la performance à un an pour une entreprise et est par nature un indicateur de long terme (pour une critique pédagogique, voir Guimbal, 2003), le publiant a besoin de temps pour publier. Il n'est sans doute pas exagéré de dire que la partie purement empirique d'un article (collecte des données) prend environ deux ans dès lors qu'il s'agit de produire des données. Il faut ensuite écrire l'article et communiquer dans des congrès. L'article s'écrit l'été pour être envoyé en traduction en octobre et une soumission à un colloque en décembre. Vous recevrez l'acceptation de votre papier en février et le présenterez en mai. Votre dernier colloque à l'*Academy of Management* en août passé, il vous restera encore quelques semaines pour tout réécrire comme on vous l'a suggéré. Si vous n'avez pas pu le faire vous perdez votre tour et donc un an. Vient ensuite le temps de la révision de la revue. Deux ou trois tours vous prendront entre un an à trois ans. Bref, il n'est pas rare d'avoir des cycles de production de 5 à 7 ans. Quand l'évaluation de votre laboratoire a lieu tous les cinq ans, cela ouvre bien des portes à la manipulation ou tout simplement à l'approximation (un laboratoire peut se voir compter un article qui aura été financé par un autre laboratoire). Il existe bien sûr des moyens de contourner le problème. Il suffit d'écrire un article utilisant des données déjà existantes (ou mieux de faire des articles sur les articles des autres, du type « j'ai étudié 100 articles dans mon champ... »). Pour réduire le cycle, le mieux est de prendre un modèle déjà bien connu et de changer marginalement une variable pour mesurer l'impact. On peut aussi interroger ses propres étudiants. Pas sûr que la science y gagne mais ce sont des pratiques courantes.

Enfin, la troisième convention est « organisationnelle ». Au-delà, d'une activité d'écriture, la recherche assemble un ensemble d'activités diversifiées que ne saisit pas la culture du comptage. *A priori*, le score d'un chercheur doit être réactualisé tous les quatre ou cinq ans. Il est donc condamné à être un publiant toute sa vie, c'est-à-dire à exercer plus ou moins les mêmes fonctions. Cette hypothèse est en contradiction avec le déroulement de la carrière du chercheur. Un bon publiant devrait passer l'agrégation et devenir professeur. Il est alors appelé à encadrer des équipes et des doctorants. Autant de tâches consommatrices de temps et qui vont le détourner de ses activités de publications. Le chercheur-encadrant pourrait alors signer les travaux de ses collègues ou de ses thésards. Cela se fait dans d'autres disciplines. Si l'habitude commence à se prendre en sciences de gestion pour la co-publication thésard/directeur de thèse, la signature collective d'une équipe de recherche avec des collègues expérimentés n'ayant lu l'article qu'en fin de cycle est une pratique à notre connaissance inexistante (ou si rare) en gestion. On le voit encore une fois, au-delà du comptage, celui-ci repose sur des conventions qui sont cette fois-ci organisationnelles. Compter ne va donc pas de soi et comprendre un chiffre signifie être capable de décrypter les *habitus* du champ organisationnel sous-jacent. Notre chercheur expérimenté pourra aussi participer à des jurys de thèse. Six ou sept par an sont les normes observées à l'heure actuelle. A raison de trois jours de travail, cette activité de soutenance lui demandera donc environ vingt jours de travail par an soit un mois complet, 1/12 de son activité. Ici en revanche, point de récompense sous forme de publication. Tout au plus l'assurance que ses collègues lui rendront la pareil en termes de temps consacré pour ses propres thésards. Plus mature, bon connaisseur des processus de publication qu'il a su dominer, notre chercheur sera peut-être enclin à résumer son œuvre dans un livre. Ce travail de maturité peut même rencontrer un

certain succès. Beaucoup de lecteurs considéreront avec sérieux cet ouvrage écrit par un collègue expérimenté dont les mots ont *a priori* plus de valeur que ceux du jeune chercheur, qui bien que (certainement) brillant, n'a pas encore passé les tests de l'institutionnalisation de son programme de recherche. En conséquence, un chercheur expérimenté devrait publier des livres, être attendu pour ses livres, plus que pour ses articles. Ces œuvres plus lourdes, plus détaillées sont dignes d'intérêt justement parce qu'il a su passer d'autres épreuves avant d'en arriver là. Le chercheur dont le livre est publié ensuite en anglais (mais aussi en Portugais pour le Brésil par ex.) accède alors à une notoriété qui rejaillit sur son institution. Il n'est d'ailleurs qu'à voir les scores h-index et g-index des chercheurs expérimentés dans Harzing pour se persuader à bon compte de l'importance des livres. Compter la performance d'un chercheur ne prend donc pas la même forme tout au long de sa carrière. Avant de mettre en place des procédures de comptage, il faut avoir défini une stratégie (encore) de ressources humaines, avec des trajectoires professionnelles bien identifiées et des objectifs différenciés à chaque étape. Le cycle de vie d'un chercheur comporte en effet différentes étapes. Si en début de carrière entre 27 et 35 ans, on peut s'attendre à une forte activité de publications, les étapes ultérieures (professorat, *tenure track*, poste de direction) doivent correspondre à un changement de métier, plus orienté vers le management et/ou la publication d'ouvrages de maturité. Or on constate aujourd'hui une superposition croissante de tâches au fur et à mesure que la carrière avance.

3.2 Les conséquences induites de l'évaluation

La prise en compte de ces conventions permettrait de mettre en place une évaluation raisonnée de l'activité de recherche. En l'état, ces conventions se retrouvent centralisées dans un seul document : le référentiel support de l'objectivation de la mesure de la performance. Un bon publiant sait où publier puisqu'il a des référentiels. Mais ces référentiels ne tombent pas du ciel. Ils sont aussi construits par des communautés. Dans ces communautés prennent place des jeux d'acteurs où les uns vont pousser leur revue préférée, d'autres celles où ils ont des chances de publier, d'autres des fantasmes avec des revues dans lesquelles aucun auteur française n'a jamais publié (ou alors un seul pour la forme). Ces classements ont l'inconvénient d'être extrêmement conservateurs. Il est difficile d'y entrer mais il sera aussi difficile d'en sortir. En effet, un bon auteur tentera d'abord de publier dans une revue classée puis ensuite se rabattra sur une non-classée. La constitution de ces référentiels occupent donc une bonne partie du temps des sociétés savantes qui vont protéger leur pré carré en essayant de faire entrer le maximum de revue de leur champ afin de faire ensuite valoir leur force. C'est un bon moyen d'avoir plus de chercheurs et au final de peser plus.

Le comptage des publications conduit à une culture de la « batonite » (ici ce sont des étoiles comme nous l'avons vu plus haut). L'important est d'être publié et cité pas forcément d'être lu et d'apporter une connaissance nouvelle. Les lecteurs croulent plutôt sous la masse des publications qu'ils ne lisent de bonnes idées émergeant de ces publications. Peut-être ce gâchis de chercheurs, ni excellents, ni mauvais pour autant, est-il le prix à payer pour voir émerger un Bourdieu par exemple. Mais dans ce cas, a-t-on vraiment progressé par rapport à la situation antérieure ? En quoi ce maelström de la mesure et ce culte de la performance permet-il de faire progresser notre compréhension du monde ? Il s'écrit aujourd'hui chaque mois plus d'articles dans nos champs de recherche que nous ne sommes capables de lire. Il faut donc faire des choix. Par exemple, ne lire que quelques revues ou quelques auteurs avec un risque de sclérose évident. Il est également possible de se spécialiser à outrance (par exemple de devenir le spécialiste de l'histoire du contrôle budgétaire en France entre 1930 et

1940) ce qui permet de ne pas être débordé par les lectures qui deviennent de fait non pertinentes. Il n'est pas sûr que cette hyper-spécialisation à outrance fasse progresser la pensée.

Cette culture de la mesure du nombre de publications conduit également à des mécanismes de gestion assez surprenants dans l'allocation des moyens financiers aux laboratoires de recherche. En effet, les budgets des centres de recherche sont actuellement affectés sur la base du score de publication des années passées. Une équipe ayant été publiante dans le passé a toutes les chances de voir ses ressources renouvelées. Si d'autres critères interviennent (gouvernance, projet de l'équipe...), il n'en reste pas moins vrai qu'au final le critère du nombre de publications l'emporte et même si une partie de l'équipe change dans les deux ou trois ans à venir. Il apparaît alors des situations paradoxales où des équipes sous performantes se retrouvent bien dotées (car elles l'ont été dans le passé) quand des équipes en constitution ou en phase de création rencontrent des difficultés de financement car elles doivent faire leur preuve. Le mécanisme de l'agrégation du supérieur en gestion rend l'exercice encore plus périlleux. Le nouvel agrégé ne choisit que modérément l'équipe qu'il va intégrer. Malheur au mal classé qui se retrouvera dans une équipe mal dotée. Il pourrait se rattacher à une autre équipe dans une autre université mais la pratique est généralement mal vue alors que ce nouvel impétrant est attendu pour aider à dynamiser l'équipe locale. Il peut aussi retarder ses publications jusqu'à sa prochaine mutation. Il arrivera alors dans sa nouvelle équipe en étant un publiant massif. De ce point de vue, la valeur du chercheur est sa capacité à publier mais aussi et surtout son stock de publication en cours de révision à un moment donné. Il sera plus confortable de muter dans une autre Université ou dans une Ecole muni d'encours de production que simplement auréolé de sa gloire passée. On imagine tous les effets pervers que ces effets de bord peuvent induire.

Une autre conséquence est liée à l'utilisation d'un référentiel unique et national. Si on peut comprendre que les instances nationales aient besoin d'un critère unique de classement, il est surprenant, dans une perspective de pilotage, que ce référentiel soit repris tel quel par les Universités et par les Ecoles (ce qui nous semble en fait moins souvent le cas pour ces dernières). Au contraire, les acteurs locaux ont besoin de référentiels locaux tenant compte de leurs propres spécificités. En effet, à quoi sert de se baser, pour certaines institutions, sur des classements pour lesquels seules les revues du bas de l'échelle sont accessibles ? Une politique de recherche se pilote sur plusieurs années en fonction de ses points forts et points faibles. Avant de viser des « revues de haut de tableau » une institution de milieu de classement aura intérêt à focaliser l'attention de ses chercheurs sur des revues accessibles. Il n'est pas non plus pertinent de considérer toutes les revues. Une Ecole spécialisée dans le commerce extérieur aura peut-être intérêt à construire une grille de référentiel qui privilégierait certaines revues en lien avec son positionnement pédagogique. En effet, à quoi lui sert de publier un article dans une revue de GRH, surtout si le contenu de cet article ne contient rien sur le management interculturel ? Déclinés localement, ces référentiels ont besoin d'exprimer une vision stratégique, garante de la pertinence de la mesure. On retrouve ainsi une prescription des recherches en contrôle de gestion. La mesure de la performance doit être articulée avec une stratégie pour être une mesure pertinente et non simplement un critère bureaucratique fondé sur une règle identique pour tous les acteurs.

Enfin, à un niveau organisationnel, ce mode d'évaluation conduit à nier une des conditions nécessaires à la vie d'un laboratoire de recherche : la diversité des rôles joués par les acteurs d'une équipe de recherche. Il est parfois étrange de voir comment le référentiel et sa philosophie gestionnaire peuvent être déclinés dans nos institutions (Hatchuel et Weil, 1992).

Doit-on par exemple considérer comme faisant partie de l'équipe de recherche, et à ce titre comme finançables, uniquement les chercheurs publiants ? Cela pose la question de jeunes chercheurs, encore non publiants, mais aussi des vieux chercheurs qui sont encore publiants mais ne participent plus à la publication et plus largement des autres membres de l'équipe qui, sans être publiants, contribuent à produire de la valeur pour l'équipe. Sorte de laborantins de la recherche, ils peuvent contribuer à faciliter la vie des chercheurs-publiants et ainsi les rendre plus efficaces. Il serait bon alors de laisser les équipes s'organiser comme elles le veulent en intégrant ou non des membres plutôt sur la base d'un projet de recherche que sur un simple critère de publiant. Il s'agirait alors de créer une sorte d'*affectio societatis* où une équipe est un projet, des liens d'intérêt mutuel entre ses membres et donc une mesure de ce projet. Il est également possible de créer des moyennes de publication par équipe et de ne plus juger des individus mais une équipe, à condition que celle-ci soit libre de s'organiser comme elle l'entend.

3.3 Neuf propositions pour une évaluation apaisée

L'évaluation est nécessaire et saine pour une bonne pratique professionnelle. Elle doit toutefois être exercée de façon apaisée en tenant compte du contexte social et organisationnel dans lequel évoluent les évalués. Cela est vrai pour les chercheurs mais également pour les autres métiers. Le problème est sans doute exacerbé chez les chercheurs dans la mesure où ils se rapprochent de ces professions où il est difficile de mesurer les résultats et où la connaissance du processus de transformation est mauvaise. Cela légitime une évaluation par les pairs (comme pour les médecins ou les magistrats) qui n'exclut toutefois pas le recours à des données chiffrées, à condition de les manipuler avec précautions. Avant toute procédure de quantification, il semble qu'il serait souhaitable de se livrer à quelques activités préparatoires.

Définition du groupe d'appartenance

- Il faut dans un premier temps **segmenter les disciplines**. Les règles institutionnelles ne sont pas les mêmes en gestion, en économie, en sociologie ou en histoire. Les langues de travail peuvent être différentes (en histoire, le français est encore reconnu). De même, le nombre de revues n'est pas forcément le même en fonction de la date d'institutionnalisation de la discipline. Le leadership exercé par les chercheurs français au plan international va conduire à s'imposer plus ou moins facilement. Si la sociologie française jouit d'une bonne réputation au plan international, ne serait-ce que pour des raisons historiques, il n'en va pas de même pour la gestion.
- Deuxièmement, il faut **définir stratégiquement à quels groupes de référence appartient l'entité** ou les membres que l'on veut évaluer et se demander si les indicateurs et les objectifs que l'on veut appliquer sont pertinents pour ce groupe. Une Ecole ou une Université jouant dans la cour européenne ne doit pas avoir les mêmes objectifs et les mêmes indicateurs que la même institution jouant une place au niveau national.
- Troisièmement, **les objectifs doivent être établis en fonction des attentes de l'institution**. A cet égard, les critères de l'AACSB par exemple sont sans doute plus réalistes que nos critères nationaux dans la mesure où ils rendent mieux compte de la variété des activités du chercheur. Au chercheur publiant est substitué un chercheur produisant. Pour une petite école, avoir quelques chercheurs académiques classiques et une participation importante du reste du corps professoral à des activités de diffusion

dans des revues professionnelles ou via des cas pédagogiques pourrait constituer un bon équilibre.

Replacer l'individu dans le groupe d'appartenance

- Quatrièmement, il faut **planifier des objectifs cohérents** avec la largeur du spectre des activités dont le chercheur a la charge. Autant qu'un référentiel de revues, un référentiel de compétences doit être établi.
- Cinquièmement, il faut **segmenter les profils des individus** en fonction du cycle de vie de la carrière du chercheur. Un jeune chercheur (même de cinquante ans) doit d'abord montrer son aptitude à publier sous la contrainte des revues quand le chercheur plus expérimenté atteindra le stade suprême de son art en publiant son ouvrage de référence.
- Sixièmement, l'évaluation du chercheur ne saurait se limiter à une quantification brute des publications. **Le contenu importe et doit être évalué** par les pairs et le grand public, dans une certaine mesure. Horngren est un auteur influent en comptabilité. Sa notoriété vient de livres pédagogiques vendus à plusieurs millions d'exemplaires, traduits en plusieurs langues. On peut déplorer une forme de pensée unique ou ne pas être d'accord avec sa présentation du sujet. Ce sont là des controverses académiquement légitimes mais il n'en reste pas moins un auteur de référence.

Elargir le périmètre d'évaluation

- Septièmement, les chercheurs en gestion s'ouvrant de plus en plus à des disciplines connexes (sociologie, philosophie, économie etc.), l'évaluation de l'activité doit **refléter cette multi-disciplinarité**. Dès lors, nous pouvons nous interroger sur l'utilité d'une liste propre à chaque discipline.
- Huitièmement, **le temps de la recherche** doit aussi être intégré à la construction de l'évaluation. Si les sciences de gestion s'intéressent aux entreprises, il est surprenant de constater que le temps de l'entreprise n'est pas forcément celui des chercheurs. Dès lors, est-il pertinent d'évaluer des laboratoires tous les 4/5 ans ? Par exemple, certaines universités américaines considèrent des durées de huit ans.
- Enfin l'évaluation de l'activité de recherche ne peut se résumer à la seule activité de travaux. **L'ensemble des facettes de cette activité doivent être appréhendées** et plus particulièrement les activités d'administration de la recherche (organisation de congrès, participation à des jurys de thèse, d'HDR etc.) ou d'implication au sein de la communauté de recherche (membres de comité de lecture, de comité de rédaction, participation à des conseils d'administration etc.)

Conclusion

L'évaluation et sa récente institutionnalisation vont à l'encontre de la tradition universitaire (Osterloh et al., 2008). Néanmoins, contrôler et piloter l'activité des chercheurs semble aujourd'hui nécessaire au regard de la massification de l'enseignement supérieur, de la volonté d'améliorer les processus de formation et de l'institutionnalisation des champs de recherche. Ainsi, il ne s'agit pas ici de remettre en cause ce mouvement mais plus simplement de l'analyser afin de proposer une évaluation apaisée des activités de recherche.

Or, l'état actuel de l'évaluation de ces activités, et notamment l'utilisation d'un référentiel unique d'évaluation, aboutit à poser un triple constat :

- L'évaluation des activités de recherche participe à un mouvement de managérialisation des activités publiques. Cette dynamique est aussi observée au sein de nombreuses structures publiques, par exemple, au sein des collectivités locales.
- L'évaluation, telle qu'elle actuellement pratiquée, n'intègre pas les conventions organisationnelles, sociales et techniques sur lesquelles repose l'activité de recherche. Dès lors, elle génère une évaluation étriquée et réductrice des activités de recherche.
- Au final, l'évaluation des activités de recherche génère des effets pervers conséquents. Le constat le plus inquiétant semble être celui de l'absence de lien entre la multiplication des publications et l'accroissement des connaissances. Les chercheurs cherchent à se jouer des règles de l'évaluation. Indirectement, leur mission première est oubliée au profit d'une seule volonté : publier pour répondre aux critères de l'évaluation.

Bibliographie

Berland ; N., Chevalier C., et Sponem S. (2008), « On ne gère bien que ce que l'on mesure », in Anne Pezet et Samuel Sponem (eds), *Petit bréviaire des idées reçues en management*, La découverte, p. 159-168.

Desrosières A. (1993), *La politique des grands nombres. Histoire de la raison statistique*, La Découverte, Paris.

Guimbal, P, (2003), « Du bon usage du retour sur investissement », *L'Expansion* 24/09/2003.

Hatchuel A. et Weil, B. (1992), *L'expert et le système*, Economica, Paris.

Hood C. (1991), "A Public Management for All Seasons", *Public Administration Review*, volume n°69, n°1, p. 3-19.

Hood, Ch. (1995), 'The New Public Management in the 1980s; Variations on a Theme', *Accounting, Organizations and Society*, Vol. 20, n° 3, pp. 93-109.

Hopwood, A. (2008), Taking the European Accounting Review forward, Letter de la European Accounting Association, mars.

Miller et O'Leary (1987) "The governable person" *Accounting Organization and Society*, Vol. 18, n°7/8.

Power, M. (1999), *The Audit Society: Rituals of Verification*, Oxford University Press, USA.

Osterloh, M., Frey, B. S., Homberg F. (2008), « Le chercheur et l'obligation de rendre des comptes », *Gérer et Comprendre*, mars, n°91, p. 48-54.