

Altres propostes i criteris avaluadors de la recerca

Alain d'Iribarne (director de Recerca del CNRS)

Francisco Llera (Catedràtic de Ciència Política Universitat del País Basc)

Alain d'Iribarne

Les enjeux de la place accordée aux revues de rang «a» dans le dispositif d'évaluation des Scientifiques Académiques Français¹

L'évaluation est de plus en plus à l'honneur tant les pressions exercées par les budgets aussi bien que par la montée en puissance des professionnels de la gestion, conduisent à interroger les décideurs sur la pertinence des usages qu'ils font des moyens qui leur sont alloués.

Les gestionnaires de la science et des scientifiques, n'échappent pas à cette dure loi, que le questionnement provienne directement du politique dans l'exercice normal de la rationalisation de ses choix budgétaires, ou qu'il provienne indirectement des citoyens souhaitant avoir leur mot à dire dans les choix scientifiques et l'usage des moyens qui leur sont alloués².

Dans un cas comme dans l'autre, l'exercice de l'évaluation ne va pas sans tensions entre les principaux acteurs collectifs concernés: les scientifiques, les gestionnaires de la science et les politiques. Cette tension s'est accentuée ces dernières années sous l'action combinée d'un appauvrissement relatif des finances publiques et d'une augmentation des moyens réclamés par les scientifiques pour faire de la «bonne science», tandis que par ailleurs, les résultats produits par la recherche devenaient des enjeux croissants de la compétitivité économique. Des «résultats» ont donc été demandés à la recherche et aux chercheurs, les gestionnaires ayant entre autres besoins de rendre «visible» ces activités pour justifier leur existence³.

Ces questions d'évaluation tendent à prendre une place d'autant plus importante et conflictuelle dans la vie du personnel de recherche que d'une part elles

1. On utilisera ici le terme de «scientifique» dans un sens générique, c'est-à-dire sans faire de distinction entre les statuts d'Enseignant-Chercheur et de «Chercheur».
2. Cette volonté est d'autant plus perceptible que l'on se trouve dans des secteurs de la recherche vis-à-vis des résultats desquels, le «citoyen» se sent directement concerné, comme celui des «sciences du vivant» qu'il est de plus invité à se financer directement à travers des collectes caritatives.
3. A moins que ce soit pour justifier leur propre existence, ce que disent volontiers les chercheurs qui ne les portent pas dans leur cœur, même si ce sont souvent des «collègues».

s’accompagnent de durcissements dans les conditions de répartition des moyens, des carrières ainsi que des honneurs, et que d’autre part elles tendent à sortir du cénacle des «pairs» pour devenir un support de médiation entre les «professionnels» de la production scientifique et leurs partenaires de toutes natures.

C’est donc dans un contexte de changement profond des modalités d’exercice de l’activité des chercheurs et de tous ceux qui contribuent au renouvellement des connaissances scientifiques et technologiques, que doivent être replacés les débats sur l’évaluation. D’un point de vue technique, ces débats concernent trois dimensions qui sont fortement interdépendantes, faisant système: la conception de l’évaluation, son instrumentation et son inscription dans des pratiques gestionnaires. Suivant leurs modalités d’agencement, ces dimensions sont, en effet, susceptibles de se contrecarrer ou au contraire de se renforcer dans leurs effets. D’un point de vue conceptuel, la situation est encore plus complexe par le fait que tout exercice d’évaluation n’a de signification que par rapport à l’acte de gestion dans lequel il s’inscrit.

L’acte d’évaluation dans le domaine scientifique tire ainsi sa complexité du fait qu’il se trouve au carrefour de deux ensembles qui doivent être analytiquement distingués tout en étant organiquement liés: le corps social des chercheurs et la connaissance scientifique. Dans le premier cas, cet acte relève de la gestion des «Ressources Humaines» et correspond à une profession particulière qui exerce ses activités dans le cadre d’institutions spécifiques dotées de leurs propres règles plus ou moins formalisées dans un statut. Dans le second cas, il relève de la gestion de la production et de la valorisation économique et sociale de connaissances et de découvertes qui sont elles mêmes plus ou moins incorporées dans des innovations⁴.

D’une façon générale, les pratiques d’évaluations mises en place ont beaucoup de mal à harmoniser les exigences d’efficacité dans la gestion de ces deux ensembles en raison des décalages structurels qui tendent à régulièrement s’établir entre eux. Il existe, en effet, pour la recherche académique quatre sous ensembles qui ont leurs logiques et leur dynamiques propres: celui des conditions à réunir pour favoriser une production féconde de connaissances nouvelles; celui du corps social des personnels exerçant des activités professionnelles de recherche; celui des institutions supports de l’organisation des activités de recherche; enfin, celui des attentes formulées par la société vis-à-vis de la recherche.

Ces décalages sont particulièrement frappants pour la recherche française qui n’arrive pas à reconstruire un système d’évaluation des chercheurs académiques capable d’harmoniser les exigences contradictoires de ces quatre sous-ensembles. Cette dysharmonie est très liée aux décalages qui se sont peu-à-peu instaurés entre les attentes aujourd’hui formulées par la société à la science et les fondements sociétaux de l’identité professionnelle des chercheurs acadé-

4. On se réfère ici d’un coté à une conception élargie du Système de production de Sciences & Technique en incluant en partie la R&D et la production d’innovation, tout en le réduisant, d’un autre coté, à sa partie publique.

miques français. Elle est cependant très variable selon les domaines d'activités scientifiques, sachant que dans tous les cas, la place accordée aux «revues de rang A» dans les procédures d'évaluation constitue un enjeu majeur de la réharmonisation. Les diverses disciplines traditionnellement constitutives des Sciences Humaines et Sociales, n'échappent pas à cette problématique

Des étapes dans l'évolution du modèle organisationnel de la science

Qu'on le veuille ou non, la production de connaissances relève d'une organisation collective qui structure son fonctionnement. Cette organisation collective des connaissances scientifiques n'est pas indépendante des modalités générales d'organisation des activités productives. On peut ainsi dire qu'à un moment donné, il tend à exister des modèles généraux d'organisation productive qui sont ensuite déclinés de façons plus ou moins différenciées suivant les pays et suivant les secteurs d'activité et, au sein du secteur d'activité de la recherche, suivant les disciplines, en fonction de spécificités «sociétales» et productives. Suivant une perspective «diffusionniste», ces modèles naissent de façon privilégiée dans une société ou un pays donné à un moment donné et dans des secteurs économiques donnés. Ils se diffusent en suite de façons plus ou moins adaptées et conflictuelles en fonction des capacités ou des volontés des acteurs concernés à se les approprier: aucun domaine d'activité scientifique échappe à cette double logique.

Ainsi, plusieurs étapes marquent le modèle organisationnel de la science en relation avec les attentes formulées par la société aux scientifiques. Elles vont du modèle artisanal au modèle de «l'économie de la connaissance» en passant par celui de la «big science».

L'artisanat de la science comme modèle nostalgique de référence

Les activités scientifiques ont ainsi un socle commun d'organisation qui révèle une forte homologie avec le modèle de l'artisanat traditionnel: dotées de peu de moyens techniques eux-mêmes le plus souvent conçus et fabriqués sur place par les chercheurs eux-mêmes à partir de leurs expériences empiriques, travaillant en solitaire ou dans des petites structures avec une faible division du travail, les chercheurs ne sont guère des professionnels de la recherche dans leur métier, ils exercent des activités qui les conduisent à chercher. La recherche académique bénéficiant de subsides publics est peu développée même dans les universités, tandis que les sociétés savantes ou les académies jouent un rôle central. Elles se trouvent en effet tant en position de magistère que de contrôle et d'évaluation des productions scientifiques à travers les communications présentées et des publications dans leurs bulletins.

Dans cette science, le chercheur revêt volontiers l'habit du savant érudit qui œuvre de façon désintéressée —si ce n'est des honneurs— pour le bien de l'humanité. Il n'a de compte à rendre à personne si ce n'est à ses pairs, le «Prince éclairé» ayant le statut reconnu de «protecteur des sciences et des let-

tres». C’est la communauté —la corporation— des savants qui définit et qui est garante des règles et des pratiques.

On perçoit bien que dans ce modèle, en dépit du caractère affiché de l’universalité de la connaissance scientifique et de la science, ainsi que de l’importance des échanges pouvant exister entre quelques grands savants, l’encre dominant de la production scientifique se fait dans des petits cénacles de proximité où souvent les réputations se font et se défont au gré des coteries et des intrigues: les investissements sociaux sont souvent aussi importants que les investissements scientifiques. La communication scientifique et le livre — l’ouvrage— jouent un rôle essentiel dans la circulation des connaissances parmi ceux qui les produisent et dans leur diffusion au sein de la société, auprès d’un public élargi mais «éclairé».

Ce modèle a largement perduré tout en s’hybridant à travers l’histoire contemporaine jusque à la veille de la deuxième guerre mondiale. Il reste très vivace en France, non seulement dans l’imaginaire populaire avec la figure emblématique du savant, mais aussi dans celle des chercheurs qui y voient un âge d’or où la «science appartenait à la science», et où en particulier, les «bureaucrates» et autres empêcheurs de tourner en rond étaient inexistants.

Il reste la référence affichée de la recherche en sciences humaines qui se voulant plus érudite que les autres appuie volontiers son évaluation sur des canons établis à travers une histoire longue et en sciences sociales qui considère plus volontiers l’exercice d’évaluation comme «non épistémologiquement fondé» et dominé par des compromis sociaux à contenu idéologique plus ou moins marqué.

L’irruption de la «big science» comme modèle dominant l’organisation scientifique dans l’après guerre

Sur ce socle initial qui est loin d’avoir totalement disparu, et dans sa continuité, est venu peu-à-peu se greffer —se surimposer?— avec la première guerre mondiale un modèle de la production scientifique qui a emprunté un certain nombre de ses attributs au modèle industriel. Une instrumentation qui s’a-lourdit en physique pour traquer l’infiniment grand et l’infiniment petit emprunte plus fréquemment ses ressources aux équipements manufacturés. Les laboratoires dont les tailles augmentent, deviennent les lieux privilégiés de recherche et s’organisent plus souvent en équipes avec des débuts de division du travail plus poussés. Au même titre que les Universités, des industries comme la chimie deviennent des lieux de recherche beaucoup plus explicites. Tout en restant fortement marquée par le modèle artisanal évolué, la recherche académique publique, au sens moderne du terme, émerge avec des politiques de l’État.

Mais on peut dire qu’une véritable rupture de modèle s’est opérée avec le modèle de la *big science* aujourd’hui bien connu, puisqu’il s’est pratiquement imposé en modèle de référence à l’ensemble de l’organisation scientifique. On rappellera simplement qu’il est né aux USA avec la deuxième guerre mondiale en Physique Nucléaire, servant de support organisationnel pour la course à

la production des connaissances nécessaires à la conception et la mise au point de la bombe atomique et plus largement des armements «modernes».

Ce modèle est lui aussi très lié à celui qui s'est mis à dominer l'organisation industrielle de l'époque à la recherche d'économies d'échelles et de gains de productivité pour des productions massifiées et plus standardisées. Il a conduit à structurer l'organisation collective de la production des connaissances scientifiques sur la base d'institutions publiques comme le CNRS en France, associé à d'autres institutions plus spécialisées par domaine scientifique. Cette organisation vient s'appuyer sur des débuts de véritables politiques scientifiques étatiques qui leur apportent des moyens humains et financiers sans commune mesure avec toutes les époques précédentes: la recherche s'établit peu-à-peu comme un secteur d'activité économique producteur de connaissances avec des personnels de recherche qui lui sont dédiés et des spécialisations internes beaucoup plus poussées. Elle devient obligée de gérer ses chercheurs tandis que ses laboratoires traditionnels, eux-mêmes plus fortement structurés, voient leurs équipes de recherche s'insérer dans des «consortiums» nationaux et internationaux de plus en plus importants, en association à l'usage d'équipements de plus en plus gros et spécialisés.

Ce modèle qui s'est amplifié dans le domaine de la physique expérimentale en passant de la physique nucléaire à la physique corpusculaire avec par exemple le développement d'accélérateurs de particules comme ceux du CERN à Genève, s'est peu-à-peu étendu aux autres domaines de production scientifique à mesure que les instruments d'expérimentation et d'observation sont devenus de plus en plus «lourds», c'est-à-dire de plus en plus coûteux. Cette extension est facilement perceptible dans des secteurs relevant par exemple des sciences de l'univers avec les observations spatiales ou des sciences de la vie avec le séquençage du Génome humain. Dans ces secteurs où dominent l'expérimentation et l'observation, l'appropriation de ce modèle industriel par les chercheurs académiques s'est d'autant mieux réalisée, qu'il est rapidement apparu incontournable compte tenu de sa supériorité productive.

Par contre, c'est principalement sous l'impulsion des gestionnaires des activités scientifiques principalement à la recherche d'économies d'échelles dans la gestion bureaucratique des institutions les concernant, que ce modèle s'est également étendu aux sciences humaines et sociales. D'une façon générale il y a été plutôt mal accueilli, au motif évoqué d'une supériorité des modes artisanaux de production dans les «sciences érudites». Cet accueil a toutefois été moins mauvais là où des besoins d'acquisitions et de traitements de matériaux empiriques étaient importants tels que l'archéologie —avec des campagnes de fouilles— pour les sciences humaines, et la sociologie ou l'économie —avec la constitution de bases de données—, pour les sciences sociales.

En Europe, cette tendance à l'accentuation de la diffusion du modèle de la *big science* dans tous les secteurs d'activité scientifique a été renforcée par l'émergence d'une politique européenne de programmes qui tend à s'appuyer sur des appels d'offres favorisant la constitution de gros consortiums européens dans le cadre de grands projets.

Une organisation scientifique sollicitée pour s’intégrer dans une économie de la connaissance

Ce modèle d’organisation de la production des connaissances scientifiques basée sur des structures lourdes, appartenant à des segments spécialisés et juxtaposés, fonctionnant en cascades, tout en poursuivant sa diffusion sur sa lancée, se voit à son tour vivement contesté en raison de sa faible capacité à traduire les éléments de connaissance issus de la science académique en innovations technologiques rapidement valorisables. Il en résulte de la part des pouvoirs publics et des industriels des demandes importantes d’inflexions du modèle au profit d’ensembles plus souples, plus maillés en réseaux dans le cadre de projets incorporant des partenaires de tous les segments de la recherche et hors recherche. L’objectif poursuivi est «d’endogénéiser» l’organisation de la recherche dans l’organisation productive élargie.

Pour simplifier, on peut dire que la transition de phase s’est amorcée en France au début des années 80 avec la reconnaissance que la crise économique initiée au milieu des années 70 n’était pas une simple crise conjoncturelle demandant des ajustements marginaux, mais représentait une crise structurelle beaucoup plus profonde. Suivant cette perspective, il a été admis que la sortie de crise impliquerait des transformations beaucoup plus radicales des logiques productives liées à des évolutions dans des normes de consommation et dans les modalités de concurrence. De même, il a été admis dès cette époque que des transformations «structurelles» seraient nécessaires dans le fonctionnement «régulateur» des institutions de recherche, que ces transformations soient ou non en relation avec une remise en cause des «économies administrées» et une affirmation quasi apologétique de la supériorité des «régulations de marchés» sur les régulations institutionnelles.

Dans ce mouvement général deux éléments peuvent être considérés comme essentiels:

- Les fondements de la valeur économique qui se sont déplacés des biens matériels vers les biens immatériels faisant de la production et de la captation de connaissances une source essentielle de la production de nouvelles richesses, et de la capacité de traduction de ces connaissances en innovations la source principale de compétitivité.
- Les modalités de l’organisation économique qui, dans une perspective de recombinaison des spécialisations insérées dans un double mouvement de décentralisation et d’intégrations fonctionnelles, se sont mises à privilégier la logique d’organisations en projets et en réseaux supportés par des technologies renouvelées d’Information et de Communication.

Il est résulté de ces deux mouvements, que les activités de productions scientifiques ont commencé à faire l’objet d’une double intégration nouvelle pour elles. Du point de vue de leurs finalités productives, elles sont devenues des parties intégrantes d’une logique fortement élargie de finalités militaires ou humanitaires, à des finalités de guerres économiques. Du point de vue de leur organi-

sation productive, elles se sont vues intégrées dans des réseaux productifs élargis à la production de connaissances nouvelles et à leur traduction en innovations principalement technologiques incorporées dans des produits ou des services renouvelés le plus rapidement possible. La recherche académique traditionnellement construite sur des bases disciplinaires «orientées sujet», se voit beaucoup plus sollicitée par de la recherche interdisciplinaire ou pluridisciplinaire «orientée objet», supportée par des consortiums mixtes recherche publique / recherche privée. De producteurs de biens publics, les chercheurs sont appelés à devenir de véritables entrepreneurs de recherche, capables de «développer» des applications de leurs découvertes, ou, à défaut, d'en négocier le développement dans le cadre de contrats éminemment marchands. La régulation par le marché à travers des contrats privés, remonte ainsi dans les activités de recherche «amont» jusque là plus habituées à travailler dans le cadre de contrats publics.

Ce nouveau modèle d'organisation productive de la recherche qui est beaucoup plus basé sur de l'intégration et de la rétroaction entre les segments spécialisés, s'est rapidement imposé dans les secteurs scientifiques dont la production alimente directement les activités industrielles, comme par exemple ceux des Sciences de l'Ingénieur qui alimentent tous les secteurs industriels intéressés par les Technologies de l'Information et de la Communication et ceux des Sciences du Vivant qui, par exemple avec le génie génétique, alimentent les secteurs pharmaceutiques et agro-alimentaires. Mais d'autres secteurs scientifiques en apparence plus éloignés tels que ceux relevant des Mathématiques et de la Physique de base avec les sciences des matériaux ou ceux relevant des sciences Physiques nucléaires et corpusculaires avec les innovations issues de la conception et des usages des instruments, se trouvent eux aussi directement concernés. Il en va également pour de nombreux secteurs des sciences humaines et sociales qui se trouvent eux aussi invités à mobiliser leurs capacités pour faciliter la traduction de leurs recherches en innovations sociales aussi bien qu'institutionnelles ou organisationnelles.

Un des traits caractéristiques de ce nouveau modèle est qu'il entraîne une perte supplémentaire d'autonomie du corps social des chercheurs tant dans la définition des orientations des activités de productions scientifiques que dans la définition de leurs temporalités. Les contraintes d'orientation et de temps, tendent en effet à remonter dans les chaînes de production de connaissances, à partir des maillons les plus proches des contraintes industrielles de production, vers les maillons constitutifs des productions de connaissances académiques traditionnelles.

Des évolutions parallèles, mais décalées dans les dispositifs d'évaluation des chercheurs

Parallèlement à ces évolutions des modèles productifs des connaissances scientifiques et de leurs modalités de mise en œuvre, les dispositifs d'évaluation des chercheurs connaissent eux aussi des aménagements substantiels. Ils suivent en quelque sorte des chemins parallèles aux outils de gestion des entreprises qui

s'instrumentalisent et s'inscrivent dans des procédures de plus en plus formalisées, tandis que par ailleurs elles glissent d'une évaluation par les moyens mobilisés vers une évaluation par les résultats obtenus. Dans un cas comme dans l'autre, les mouvements s'opèrent sous l'influence des pratiques anglo-saxonnes qui tendent à s'instaurer en normes internationales, non sans rencontrer de solides réticences dans de nombreuses disciplines du secteur des SHS.

L'évaluation comme compromis social: la critique de l'évaluation par «les pairs de proximité»

En accompagnement avec la mise en place de nouvelles institutions de recherche dans l'après guerre, l'évaluation des activités de recherche et des personnels qui y participent s'est faite au sein de commissions ou comités constitués par des «pairs» qui en principe fonctionnaient comme les garants de l'exercice de ces activités aux canons des règles de l'art des pratiques de la sciences. Dans ces commissions, la désignation de plusieurs rapporteurs, le caractère contradictoire des débats, l'explicitation des bases de l'évaluation ainsi que la prise en compte des conditions de contextualisation, assurent en principe subtilité et équité dans les jugements, ce qui doit être à même de satisfaire ceux qui sont soumis à ces procédures.

En fait la subtilité des éléments retenus pour différencier des dossiers, le caractère souvent opposé des arguments avancés en faveur ou au détriment d'un dossier à un moment ou un autre en fonction des circonstances, ainsi que l'évocation d'éléments extérieurs aux dossiers, doublés de difficultés à convenablement contrôler le bon fonctionnement des procédures de vote, a souvent conduit à faire des reproches aux procédures d'évaluation. Il s'agit le plus souvent des reproches de trop grande subjectivité ou de trop grand relativisme dans les processus d'évaluation qui ne sont pas uniquement le fait des évalués déçus.

Ils correspondent malheureusement à des constats assez fréquents de partialité dans les jugements, de dominations exercées par certaines personnes ou certains groupes, ou d'arrangements et marchandages effectués entre personnes ou groupes qui se comportent comme autant de portes paroles d'intérêts particuliers au détriment des intérêts généraux et en tout les cas, de l'équité. Ainsi les résultats des processus d'évaluation tendent à apparaître comme le fruit de compromis trop peu transparents, trop ancrés dans des compromis locaux et finalement peu satisfaisants au regard des objectifs fixés, ce qui risque de les discréditer. C'est pourquoi, les chercheurs eux-mêmes ont été conduits à souhaiter un meilleur encadrement des pratiques de compromis en s'appuyant sur des procédures plus formalisées et sur un outillage plus externe aux acteurs directement impliqués dans les procédures

Le besoin ressenti d'une «rationalisation» des procédures d'évaluation

Ces questions d'égalité des candidats face aux procédures, de recherche d'équité et de transparence dans les décisions prises, couplées à des besoins de

plus en plus ressentis de pouvoir établir des comparaisons entre des dossiers de candidatures ou entre des résultats de processus d'évaluation, ont donc poussé à rechercher des dispositifs qui permettent «d'encadrer» les procédures à défaut de pleinement les «objectiver». Le problème était en quelque sorte de trouver des éléments de jugement qui échappent à la subjectivité des évaluateurs de proximité et qui leur étant extérieurs s'imposeraient à eux, constituant ainsi des critères objectifs.

Ce besoin de rationalisation dans l'évaluation des activités des personnels de recherche, s'est par ailleurs d'autant plus fait ressentir que, dans le cadre des procédures de recrutement et de promotion généralement formalisées et collectives, les charges de travail des commissions se sont beaucoup alourdies en raison de l'augmentation considérable du nombre de personnes à évaluer. A joué dans le même sens, un accroissement des sélectivités tant pour les recrutements que pour les promotions, sélectivités elles mêmes liées à la dégradation des ratios du nombre de postes par candidat.

Enfin, au-delà des besoins de rationalisation, les besoins de normalisation à des standards internationaux se sont fait sentir en accompagnement de l'internationalisation de la gestion des scientifiques, l'émergence de ces pratiques d'évaluation et leur tendance à se diffuser dans les diverses disciplines y compris dans les sciences sociales et dans les sciences humaines, en imitation des pratiques anglo-saxonnes, conduisant en effet à l'établissement de véritables normes d'évaluation internationales qui tendent peu à peu à faire référence. De telles normes permettent aux évaluateurs d'étalonner de façon comparative les résultats de leur soutien aux divers domaines d'activités scientifiques à partir de standards autorisant des comparaisons internationales. Il permet à l'inverse aux évalués de se situer par rapport à la concurrence élargie ou par rapport à des perspectives d'activités de recherche beaucoup plus ouvertes sur l'international en raison d'une certaine normalisation des référents utilisés. Pour les chercheurs européens, ces évolutions tendent à accompagner l'émergence dans un espace de recherche européen d'un véritable marché du travail européen des scientifiques. Elles tendent donc à favoriser une internationalisation de la production scientifique

C'est dans cette perspective que s'inscrit l'apparition dans les processus d'évaluation des activités de recherche, des tentatives pour mesurer les résultats de ces activités en s'appuyant sur la constitution d'indicateurs quantifiés. Plus largement, l'apparition de ces indicateurs correspond à la mise en œuvre d'outils non seulement destinés à faciliter l'évaluation en la «normalisant» à partir de «standards», mais éventuellement à l'industrialiser en instituant des procédures de traitements automatisés des données, à plus ou moins grande échelle.

La publication constituant l'aboutissement traditionnel de toute activité de recherche et la base du jugement des pairs sur la qualité du travail réalisé, il était en quelque sorte naturel que les publications gardent un statut particulier dans l'évaluation du travail des chercheurs et soient, de ce fait, le support privilégié de tentatives de codification.

Il en est résulté l’apparition dans le paysage de l’évaluation, d’une partition hiérarchisée entre les publications qui a conduit à faire émerger une catégorie singulière de publications qualifiées de «revues à comités de lectures». Cette catégorie a elle-même tendu à être plus précisément spécifiée avec la mise en exergue de revues dites de «rang A». Une autre étape a également été franchie dans la recherche d’une instrumentation d’évaluation performante et objective en utilisant des «citation index» en corollaire de la scientométrie. Dans ce cas, l’évaluation n’est plus le volume de publication, mais le volume de citation par les pairs. Ce volume est sensé traduire un indice qualitatif d’importance de contribution à la connaissance scientifique.

Ainsi, plus ou moins consciemment, les scientifiques confrontés à une gestion d’une «science de masse» ont-ils introduit dans leurs pratiques des méthodes d’évaluation de masse. Tout à fait logiquement, ces méthodes sont apparues en premier dans les secteurs scientifiques où elles correspondaient aux ruptures les plus faibles, c’est-à-dire dans les sciences de la nature, là où les articles constituaient depuis toujours la seule base légitime de la production scientifique et là où le statut des revues était fortement différencié et internationalement reconnu.

La problématique sous-jacente à cette démarche, revient, pour les évaluateurs de proximité à déléguer à un «tiers certificateur» le rôle d’évaluer la qualité du travail scientifique: les pairs éloignés —«referee» et membres des comités éditoriaux— dans le cas de l’évaluation par des revues; la communauté mondiale des pairs dans le cas des index de citation. Toutefois, il faut souligner que dans le cas d’une utilisation des revues de «rang A», le rôle des pairs de proximité ne se limite pas à décompter des quantités d’articles produits. Il consiste surtout à spécifier des profils de référence à partir du choix du «mix-revues» retenu pour évaluer. En effet, le principal problème qui est posé aux évaluateurs est de déterminer la liste des revues de «rang A» qui serviront de bases à leur évaluation. Là aussi, la détermination de cette liste peut être soit «externalisée» en la déléguant à une institution extérieure —par exemple les économistes «purs et durs» tendent à se référer à la liste établie par le JEL, *Journal of Economics Literature*— soit à rester endogènes à la procédure de proximité, cette liste étant spécifiée par les évaluateurs eux-mêmes en établissant des références mixtes à partir de diverses listes, par exemple en fonction de leur degré d’ouverture à des perspectives pluridisciplinaires. D’une certaine façon, l’usage de «citation index» permet de renforcer la sélectivité des évaluateurs puisque dans ce cas, ils se réfèrent non plus à des quantités de publications mais à des scores de citations par les pairs qui correspondent à des indicateurs beaucoup plus polarisés.

Par rapport à l’évaluation artisanale de proximité initialement utilisée, cette introduction dans les procédures de critères plus formalisés, standardisés et externalisés constitue indubitablement un progrès en assainissant les bases de l’évaluation. Toutefois elle ne manque pas, dans le domaine des SHS, de critiques virulentes et pas toujours injustifiées. En effet, les chercheurs français ne manquent pas de dénoncer tour à tour: sous couvert d’internationalisation,

la domination des revues anglo-saxonnes et leur incapacité à accepter le *French style* de la recherche et de l'écriture; le «verrouillage» des comités de revue et la partialité des «referee» en relation avec des écoles ou des chapelles; l'influence dans les scores de citation, des notoriétés et des «réseaux» participant à des auto-célébrations ou à des auto-contrôles.

Les manques de congruence entre les critères d'évaluation des scientifiques et les attentes formulées à la recherche par la société

Ainsi, à travers ces mouvements qui interviennent dans la façon dont les sociétés sollicitent l'organisation scientifique en fonction des finalités qu'elles lui assignent, la recherche académique traditionnellement produite de façon dominante dans ces institutions publiques de formation et de recherche que sont les universités ou diverses autres institutions publiques plus exclusivement consacrées à la recherche, se voit régulièrement questionnée sur ses pratiques. Il lui est demandé de bouger, l'acteur principal de ces remises en cause étant les pouvoirs politiques, eux-mêmes poussés par diverses composantes de la société, en particulier les acteurs économiques et les citoyens.

En parallèle, en grande partie sous l'impulsion des chercheurs gestionnaires des chercheurs, les pratiques d'évaluation des chercheurs par les institutions de recherche, évoluent également, l'objectif étant de sélectionner et de gratifier ceux qui à leurs yeux sont les «meilleurs chercheurs». Dans ce mouvement, l'évaluation par des «revues de rang A» ou la «citation index» tendent à prendre une place centrale à des degrés divers selon les traditions des secteurs scientifiques ainsi que selon les constructions sociétales des identités professionnelles du corps social des chercheurs.

Le problème qui se pose aujourd'hui aux décideurs publics chargés de la gestion de la recherche est de savoir dans quelle mesure la place centrale accordée à ces indicateurs dans un référentiel d'évaluation, favorise ou non une inscription de la recherche académique dans une économie du savoir et plus largement dans des sociétés du savoir dont l'objectif serait de disposer de savoirs scientifiques récents et largement diffusés et appropriés dans l'ensemble de la population. Pour tenter de répondre à cette question on examinera successivement les effets potentiellement pervers que peut avoir une évaluation des chercheurs trop exclusivement centrée sur les revues de rang A ou sur des index de citation, au regard d'une part d'un panel plus large de publications et d'autre part de critères élargis à d'autres types de productions.

Les effets pervers potentiels d'une restriction des publications aux seules revues de rang A

Il est intéressant d'examiner les conséquences potentielles que peut avoir une domination de l'évaluation des chercheurs par des revues de rang A sur la fécondité de la recherche, elle-même évaluée en terme de renouvellement des connaissances scientifiques. Ces conséquences peuvent être évaluées au regard

de restrictions dans les choix des sujets et objets de recherche; de prime à la «science normale» et de limitation de l'interdisciplinarité; d'allongement des délais de publications. Les effets potentiellement les plus pervers dépendent beaucoup de la pluralité des revues et de l'ouverture des évaluateurs dans la constitution des listes de revue.

En matière de choix des sujets de recherche, il existe un risque important de limitation des sujets et surtout d'apparition de décalages avec les besoins de connaissances des autres acteurs de la société, dans la mesure où les instruments de contact avec ces autres acteurs sont fortement atrophiés, les supports se limitant à des outils qui sont quasi exclusivement réservés à une «communication de la science pour la science». On sait en effet que les questions posées aux scientifiques par d'autres qu'eux-mêmes peuvent constituer des questions qui convenablement retraduites en questions de recherche, peuvent s'avérer particulièrement fécondes. Ces risques peuvent cependant être limités s'il existe par ailleurs des lieux et des moments de contacts réels et approfondis entre les chercheurs et «les autres».

Les risques de standardisation de la pensée et des pratiques de recherche relèvent d'autres mécanismes et d'autres logiques. Les revues dominantes sont généralement des revues disciplinaires ancrées dans les *mainstream* des disciplines, avec pour effet de fortement favoriser des approfondissements théoriques donc de favoriser la production et la reproduction d'une «science normale», au détriment d'un élargissement vers d'autres approches dont les fondements théoriques sont généralement moins assurés parce que s'inscrivant dans ces «bricolages» qui caractérisent de façon dominante les connaissances en cours d'émergence. Toutefois les jugements que l'on peut porter sur les effets de ces dominations sur la fécondité de la science dépendent beaucoup de la vision que l'on a des mécanismes de la production des connaissances scientifiques, et en particulier de la place que l'on accorde aux notions de paradigmes scientifiques, ainsi qu'aux conditions d'émergence et de renouvellement de ces derniers: les dommages en la matière seront jugés d'autant plus importants que l'on aura une vision de la progression des connaissances scientifiques qui s'appuiera de façon privilégiée aux marges des disciplines et sur une dynamique des disciplines appuyées sur des recompositions par apports externes.

En fait on voit bien que la question soulevée ici est moins celle de la domination exercée par une revue ou une autre, que celle de la pluralité plus ou moins grande des revues et du verrouillage de leurs conditions d'accès. Si on admet en effet que la fécondité de la science dépend majoritairement de sa capacité à favoriser l'émergence et la diffusion de conceptualisations alternatives et concurrentes, alors ces questions de pluralité dans les revues de rang A deviennent essentielles. Cela pose en particulier le problème du classement relatif des revues à caractère interdisciplinaire et plus spécialisées sur un domaine d'application, par opposition aux revues disciplinaires et non spécialisées dans un domaine. Cela pose en contrepoint deux nouveaux problèmes: celui de la place accordée aux recherches plus appliquées par rapport aux recherches

plus théoriques; celui des délais et des moyens nécessaires pour qu'une revue nouvelle puisse émerger et être légitimée dans un classement. On voit bien que la question posée en corollaire à la pluralité et à l'ouverture des revues est celle de la pluralité et de l'éclectisme des revues qui seront retenues par les évaluateurs de proximité pour constituer leurs référentiels d'évaluation.

Un autre effets pervers de la concentration de l'évaluation sur des revues de rang A, est qu'il en résulte un afflux de propositions d'articles à ces revues, avec comme conséquence la création de files d'attentes et de délais toujours croissants entre le moment où la recherche se fait et le moment où ses résultats sont enfin diffusés. De ce point du vue il est particulièrement intéressant de constater que ce sont les scientifiques des secteurs où ces mécanismes d'évaluation sont les plus dominants — ceux de la physique et des sciences de la vie — qui les premiers se sont saisis d'Internet pour corriger cet effet pervers, en y diffusant leurs publications précoces dans le prolongement des pratiques antérieures de publication que constituaient les «preprint». Il en résulte aujourd'hui dans ces secteurs, une véritable dualisation du marché des publication qui se segmente entre publications rapides et non certifiées et publications lentes et certifiées. Cette dualisation vient bouleverser les pratiques de diffusion de la connaissance et tend à donner aux revues de rang A un rôle dominant de support d'évaluation des chercheurs à travers leurs actions de certification, au détriment de leur rôle historique de diffusion des connaissances scientifiques.

Si les perversités potentielles de la domination des revues de rang A dans l'évaluation des chercheurs peuvent être en quelque sorte auto limitées en ce qui concerne les précédents aspects il en va différemment pour tous les autres aspects qui intéressent les publications: les autres revues, les ouvrages collectifs et les livres. On peut mesurer ici les effet dévastateurs de cette domination, liés aux simples effet induits de la logique de pilotage: les chercheurs rationnels et préoccupés par leur réussite professionnelle sont tout logiquement appelés à désertter ces supports au profit de ceux qui ont perdu toute ambition, soit par renoncement pour les jeunes, soit par achèvement de leur carrière pour les plus anciens. Or, ces publications constituent souvent des véhicules essentiels de la diffusion des connaissances, aussi bien de publics plus localisés et qu'au contraire d'un public élargi à travers la vulgarisation scientifique. En fait, à travers ces aspects, c'est la question des ancrages et de la présence des chercheurs dans leurs territoires de proximité qui est posée à un moment où l'importance d'une présence active de leur part pour dynamiser des territoires est mis en exergue à travers la notion de «systèmes locaux d'innovations». C'est également la question de la constitution et de la diffusion d'une culture moderne qui est posée.

L'existence d'un hiatus entre une évaluation dominée par des revues de rang A et les attentes formulées par la société à la science

Ces analyses des perversités potentielles d'une focalisation de l'évaluation des chercheurs sur des revues de rang A peuvent être poursuivies dans le prolongement de ce qui précède en accordant une attention aux autres dimensions

des produits de l'activité scientifique suivant les logiques qui dominent dans les analyses de la «triple hélice» ou dans celles des «systèmes nationaux d'innovation». Dans ces analyses qui servent aujourd'hui de support aux pouvoirs publics à tous les niveaux pour orienter leurs politiques de soutien aux activités de S&T et de R&D, les publications ne deviennent qu'une composante pas toujours jugée essentielle aux cotés d'autres produits de la recherche qui viennent supporter divers chaînes de «production de valeur économique»: la valorisation économique directe auprès des entreprises à travers les dépôts de brevet et les activités de conseil, correspondant à la circulation des nouvelles connaissances cristallisées dans des «objets techniques»; la formation, support codifié mais largement immatériel de transmission des connaissances nouvellement produites; la mobilité des chercheurs en dehors des activités de recherche, support privilégié de circulation des connaissances en cours d'élaboration est non encore codifiable.

Un examen approfondi de ces nouvelles dimensions nous amènerait trop loin tant elles constituent en elles mêmes les bases renouvelées d'une problématique de la recherche et de sa place dans les activités économiques. C'est cependant dans cette perspective qu'apparaît le maximum de hiatus entre la vision d'activités de recherche pilotées par des revues de rang A, et la vision des activités de recherche tel qu'il est attendu qu'elles fonctionnent à partir d'un référent d'évaluation ainsi élargi, le problème posé étant celui de la pondération des composantes de ces référents en fonction des secteurs scientifiques, de la place dans les maillons de la recherche, ou des moments dans un processus de recherche aussi bien que d'une carrière. Il y a là tout un travail qui reste à faire pour disposer des connaissances voulues, travail qui intéresse directement la gestion de la science, ou si on préfère, son gouvernement.

On a évoqué à plusieurs reprises l'idée que ces hiatus étaient plus ou moins forts suivant les secteurs scientifiques et suivant les référents sociétaux des communautés scientifiques. Il est certain qu'il est impossible de comprendre sans cela, la force des tensions qui existent en France depuis le début des années 90, autour de la réforme de ce dispositif institutionnel d'évaluation de la recherche, qu'est le Comité National de la Recherche Scientifique, chargé entre autre d'évaluer les chercheurs et les laboratoires du CNRS. Il est également possible de comprendre pourquoi c'est globalement dans les communautés de chercheurs exerçant leurs activités dans les secteurs des sciences humaines et des sciences sociales, que ces tensions sont les plus fortes de façons récurrentes.

Conclusions: La nécessité d'instaurer un référentiel multiple pour évaluer les chercheurs

Quels dispositifs d'évaluation des chercheurs faudrait-il donc mettre en place pour avoir une science adaptée aux attentes qui lui sont aujourd'hui formulées par la société? Cette question met l'évaluation des chercheurs au cœur des dispositifs de régulation de la science à la fois dans son fonctionnement interne et dans ses relations avec la société.

Il apparaît clairement de l'expérience acquise, que, quels que soient les efforts qui peuvent être faits, l'évaluation des chercheurs reste un exercice par nature entaché d'arbitraire et qui, in fine, relève par essence du compromis social. Loin d'être une tare cette situation est une bonne chose car elle oblige à maintenir des procédures de proximité à fortes composantes humaines. Mais c'est aussi pourquoi cette évaluation a besoin d'être instrumentée et procédurée avec rigueur, car c'est cette rigueur qui lui assure les minimums de transparence et d'équité dont elle a absolument besoin pour être socialement acceptable. Elle a en particulier besoin que des règles soient établies, rendues publiques et respectées par les évaluateurs. Elle a également besoin d'être considérée comme légitime par les acteurs du système. Pour cela elle doit être efficace dans l'articulation de la gestion des scientifiques et la gestion des connaissances.

Dans cette perspective, l'évaluation a besoin d'être conçue et de s'instrumenter sans dogmatisme pour éviter l'enfermement dans une «science normale» qui conduit inévitablement à la sclérose. Elle a besoin d'être plurielle pour tenir compte des segments du système de S&T auquel elle se réfère et du moment où elle s'inscrit dans le cycle de production de la connaissance. Ainsi apparaît la nécessité d'adopter des référentiels certes formalisés, mais multiples, élargis, reconnaissant plusieurs métiers de la recherche, admettant des diversités de pratiques suivant les secteurs scientifiques et laissant ouvertes les pratiques dans chacun d'eux. En fait, le seul problème vraiment intéressant est celui de savoir dans quelle mesure les processus et procédures d'évaluation des scientifiques concourent à la fécondité de la production scientifique et à l'avancée des connaissances utiles à l'humanité.

La question ne se pose cependant que si l'on admet qu'il existe une corrélation significative entre les comportements de recherche des chercheurs et les pratiques de leur évaluation. Or, on rappellera que cette hypothèse fait débat, renvoyant à des interrogations sur la capacité des systèmes d'évaluation à dicter leurs pratiques aux chercheurs, ou plus modestement à les orienter. Les options en la matière diffèrent largement. Elles vont du sentiment de leurs faibles effets inductifs qui seraient liés au caractère hasardeux de la découverte et à la capacité des chercheurs à «subvertir» les règles et leurs applications au profit de leurs propres orientations, à des sentiments de surdétermination de leurs effets qui seraient liées à des comportements «opportunistes» des chercheurs à la recherche de gratifications.

Éléments de bibliographie

- AMABLE, B.; BARRE, R.; BOYER, R. (1996). *Les systèmes d'innovation à l'ère de la globalisation*, Paris, Economica, 401 p.
- BOYER, R.; DIDIER, M. (1998). *Innovation et croissance*, Paris, Conseil d'Analyse Economique, La documentation Française, 196 p.
- CALLON, M.; MAURICE, M.; MUSSELIN, CH. (1996). «Introduction», *Revue Sociologie du Travail*, n° 3/96, p. 253-261.

- CNRS (2000). *Un siècle de recherche technologique. Histoires croisées des sciences et des techniques*, Paris, CNRS, 237 p.
- BRANCIARD, A. (1999). *Espace d'innovation dans la biologie et recomposition d'espaces productifs*, Aix-en-Provence, LEST/CNRS, 151 p.
- CARACOSTAS, P.; MUL DUR, U. (1997). *La Société, Ultime Frontière. Une vision européenne des politiques de recherche et d'innovation pour le XXI^{ème} siècle*, Bruxelles, Commission Européenne, DG XII, 211 p.
- COLLOQUE DE LA VILETTE (1999). *L'opinion publique face aux plantes transgéniques*, Paris, Albin Michel, 213 p.
- COMBES, P.-PH.; LINNEMER, L. *La publication d'articles de recherche en économie en France, une approche par la qualité des revues*, Paris, CREST-LEI, 34 p.
- CRANNEY, J. (1996). *INRA 50 ans d'un organisme de recherche*, Paris, INRA Editions, 526 p.
- DE LA VEGA, J. (2000). *La communication scientifique à l'épreuve d'Internet, l'émergence d'un nouveau modèle*, Villeurbanne, Presse de l'enssib, 252 p.
- DOGAN, M.; PAHRE, R. (1991). *L'innovation dans les sciences sociales. La marginalité créatrice*, Paris, Puf, 322 p.
- ETZKOWITZ, H. (1994). «Academic-Industry Relations: A Sociological Paradigm of Economic Development», in LEYDESDORFF L., VAN DEN BESSELAAR, P. (ed.), *Evolutionary Economics and Chaos Theory*, New-York, St Martin's Press, p. 139-150.
- FERNE, G. (dir.) (1993). *Science, pouvoir et argent. La recherche entre marché et politique*, Paris, Éditions Autrement, 217 p.
- FORAY, D. *L'économie de la connaissance*, Paris, La Découverte, 120 p.
- GADILLE, M.; IRIBARNE, A. D'; LANCIANO-MORANDAT, C. (1998). *The French science and technical system between societal construction and sectorial specificities*, Communication au 14th Seminar of the European Group of Organisation Studies, Maastrich, 9/11 juillet.
- GADILLE, M., IRIBARNE, A. D' (2000). «L'IN2P3: une création institutionnelle originale face à la production et à la valorisation des connaissances scientifiques et technologiques», in TALLARD, M.; THERET, B.; URI, D. (dir.), *Innovations institutionnelles et territoires*, Paris, L'Harmattan, p. 205-238.
- IRIBARNE, A. D' (1996). «L'organisation et le fonctionnement de l'IN2P3», in YOCCOZ, J. (dir.). *25 ans de recherche à l'IN2P3, la science, les structures, les hommes*, Gif sur Yvette, Editions Frontières, p. 205-257.
- (1998). *La connaissance, le scientifique, le prince, le producteur et le citoyen*, Contribution au Colloque européen «L'inscription sociale de la science» organisé à Paris par la Commission Nationale Française pour l'UNESCO, les 5 et 6 novembre.
- (1999). «Le CNRS, proposition d'analyse pour une réforme», *Annales des Mines, Gérer et comprendre*, n° 57, septembre, p. 58-80.
- LANCIANO-MORANDAT, C. (1999). *Le mode de structuration du département des «sciences chimiques» du CNRS*, Rapport de recherche, Aix-en-Provence, LEST-CNRS, 161 p.
- LATOUR, B. (1995). *Le métier de chercheur au regard d'un anthropologue*, Paris, INRA Éditions, 95 p.
- LORINO, PH. (1998). «Organisation et innovation: l'organisation à la française», *Annales de Mines, Réalités Industrielles*, novembre p. 27-31.
- LUNDVALL, B.-A. (éd.) (1992). *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*, London, Pinter.

- LURCAT, F. (1995). *L'autorité de la science*, Paris, Cerf, 251 p.
- MACHAN, CH. (1988). *La nature sans foi ni loi. Les grands thèmes de la physique du XX^e siècle*, Paris, Belfond/sciences, 220 p.
- MUSTAR, PH. (1998). «Les transformations du système de recherche français dans les années quatre-vingt», *Annales des Mines, Réalités Industrielles*, février, p. 16-21.
- PESTRE, D. (1995). «De la définition des pratiques physiennes durant la guerre froide», *Revue L'Aventure humaine*, n° 2, L'américanisation de la recherche, juin, p. 11-22.
- PICARD, J.-F. (1990). *La république des savants. La recherche française et le CNRS*, Paris, Flammarion, 339 p.
- QUERE, M. (coord.) (1998). *Les technopoles en Europe*, Rennes Atalante, 272 p.
- RAMUNNI, G. (1995). *Les sciences pour l'ingénieur. Histoire d'un rendez-vous des sciences et de la société*. Paris, CNRS Éditions, 150 p.
- RESCHER, N. (1993). *Le progrès scientifique*, Paris, Puf, 37 p.
- SHAPING, S. (1998). *La révolution scientifique*, Paris, Flammarion, 257 p.
- VILKAS, CH. (2001). *L'art de gouverner la science dans le système public français*, Thèse de Doctorat de l'Institut d'Etudes Politiques de Paris, 640 p.
- VINCK, D. (1995). *Sociologie des sciences*, Paris, Armand Colin, 292 p.

Francisco Llera

Bueno, muchas gracias por la invitación, por la satisfacción de volver a encontrarme con viejos amigos, y algunos nuevos también. Y tengo que pedir disculpas porque, la verdad, es que no sabía muy bien de qué iba a hablar, quería insistir sobre esta cuestión. Y entonces hablando un poco de la primera idea, que me la dio Joan Botella, y no estaba muy bien definida... Luego vi un poco el programa y sobre «otros criterios» se pueden decir tantas cosas... Yo voy a hablar desde mi experiencia, la más reciente, que es la siguiente: haber sido coordinador fugaz pero intenso en el MEC durante el año pasado, que ya sabéis que fue un año extraordinario por la presión que tuvimos durante tres meses para evaluar y resolver todos los programas nacionales en condiciones realmente alucinantes. Por lo tanto, permitidme que sea bastante escéptico sobre esto de la evaluación de partida. Yo estaba bastante alucinado sobre el informe de las revistas que he manejado, y bastante alucinado con la parte que me ha tocado, seguro que es muy parcial porque era muy de última hora y he tenido todo el día de ayer para discutir otras cosas sobre las revistas, los criterios, en fin.

Yo me planteo realmente qué armas se deben utilizar en la evaluación, qué criterios a utilizar. Desde luego, yo creo que tenemos que ser mucho más cualitativos en la evaluación, muchísimo más discriminantes. Es decir, no hacer un criterio de evaluación genérico, ni siquiera no digo yo para las disciplinas, ni siquiera dentro de una disciplina, para todos los campos de una misma disciplina. Esto es bastante obvio en la medida que lo conocemos y sabemos que las características que tienen es la misma fragmentación. Y luego, también tengo la experiencia de haber trabajado como asesor para los planes de la ciencia, tanto españoles como del Gobierno vasco, y estar en la comisión de investigación de mi universidad, la cual como sabéis, a diferencia de Cataluña, es la única universidad pública del País Vasco. Por lo tanto, es la que absorbe la inmensa mayoría, más del 90 %, de los dineros de investigación públicos en Euskadi. Y luego, también he vivido la experiencia más reciente de haber recibido el encargo del Gobierno vasco para hacer una evaluación del rendimiento de los recursos científicos en nuestra universidad, insisto que es el 90 % de los recursos científicos en el País Vasco, tanto en Ciencias Sociales, informe que terminamos ya el año pasado, como en Ciencias Humanas. Como os podéis imaginar, los dos informes son bastante distintos. Voy a empezar por aportar una serie de datos... Creo que tenemos una cierta experiencia en el País Vasco en la cuestión de programas de investigación, quizá por el hecho de que los interlocutores sean menos. Es decir, una universidad, un gobierno, a pesar de las dificultades de relación dado el país que tenemos, pero será más fácil me imagino que un gobierno, con menos recursos que el nuestro, y además un montón de interlocutores universitarios que han aparecido poniéndose de acuerdo entre ellos. Una universidad, por otra parte, muy joven, y por lo tanto muy voluntariosa, y a veces esto también es una ventaja. La ciencia también requiere mucha experiencia, pero a veces demasiada experiencia muy buro-

cratizada puede ser un impedimento para adaptarse a la capacidad de cambio en los modelos, en la propia evaluación y en los modelos de gestión de los recursos. Por lo tanto, yo creo que es interesante ver la evolución de los datos en este terreno, específicamente en nuestro campo, en las ciencias sociales, y tener en cuenta algunos datos de opinión de los propios investigadores sobre la tendencia en evaluación y como ven ellos la tendencia en los últimos tiempos. Sobre los últimos cinco años, que yo creo que son extremadamente importantes, he llegado a la conclusión que para ver la evaluación desde la perspectiva cualitativa hay que introducir, obviamente, criterios relativamente objetivables y sobre los que tiene que haber un consenso. Éste es un elemento importantísimo en la diferenciación de los grupos. Es decir, no podemos aplicar el mismo criterio de evaluación a todos los grupos, incluso dentro de las mismas disciplinas, incluso dentro del mismo campo de especialización de la misma disciplina, porque no es lo mismo un grupo emergente que un grupo consolidado o que un grupo de alto rendimiento. Hay razones de tiempo, hay razones de carrera, hay razones incluso de líneas más o menos consolidadas y hay razones también de productividad. Y si aplicamos el criterio de alto rendimiento a todos los grupos, evidentemente, los jóvenes investigadores no saldrán nunca de la economía sumergida.

Otro aspecto importante lo constituyen, no los dineros públicos, sino los dineros privados, el mercado. Porque el mercado se compone de los dineros públicos, pero no estamos hablando de eso, desde el punto de vista de la evaluación de los rendimientos o de la aplicación o de las líneas estratégicas que, en la elaboración de los planes están teniendo en cuenta lo que el mercado requiere, sino el mercado como actor directamente implicado en la propia financiación y en la propia captación de recursos. Esto es importante y en nuestro caso, por ejemplo en el último acuerdo que ha habido entre la universidad y el Gobierno para hacer un plan universitario, hay una parte de los recursos de investigación que el Gobierno ya no sólo los va a hacer competitivos, sino que únicamente se los va a dar a la universidad. Y la universidad los va a administrar a su antojo pensando únicamente en el alto rendimiento, en los grupos más consolidados y en los de máximo nivel, y un criterio fundamental de evaluación será: si usted consigue recursos del mercado, usted mantendrá su estatus y obtendrá toda la financiación que quiera, porque tiene un plan de investigación, pero en la medida en que usted capte la mitad de los recursos del mercado. Porque si tiene un rendimiento alto, una experiencia elevada, un plan de investigación, muchas líneas de investigación, usted tendrá que tener capacidad para obtener recursos. Es el modelo de la financiación electoral alemana, es decir, yo le financo si usted consigue la mitad de los recursos en el sector privado. Éste es el planteamiento que está introduciendo nuestra universidad de acuerdo con el Gobierno.

Les daré unos datos para ver cómo se han movido los recursos. Representan una visión panorámica de cual ha sido la evolución en millones de pesetas de los dineros públicos en nuestra universidad, que, insisto, es más del 90 % de la investigación en el País Vasco, en los diferentes campos. Supone un 19 %

de los recursos en los cinco años de maduración de nuestro modelo público de investigación. Del 42 % se ha pasado al 24 % como primer punto de referencia de los recursos. El año 1998, que es el último año de referencia en evaluación, nos situamos en 1.100 millones de pesetas en un país de 2 millones de habitantes y una universidad con 3.000 profesores. Esos somos todos los investigadores, porque en Euskadi no hay CSIC ni nada parecido. En Euskadi sólo hay una universidad pública, luego hay universidad privada, pero la investigación que se realiza allí es mínima. Vosotros que tenéis vuestros datos en la mente, los de Cataluña, podéis hacer la comparación. Aquí tenemos los datos de proyectos e inversión en ciencias sociales nada más y la procedencia de los recursos. Aquí hay un dato que luego veremos, que es cuando la procedencia local es abundante, y sobre todo no se diferencia entre el grado de consolidación o de emergencia de los grupos, pues la tendencia será la de no ir a otros sitios a buscar recursos, puesto que ya los tenemos en casa. Y éste es uno de los resultados que está cambiando la estrategia del propio Gobierno, de las universidades, de los criterios de evaluación, justamente para evitar este riesgo de fragmentación y localismo en la captación de recursos y, por lo tanto, en cierta medida, en la competitividad. Fijaos que el 67 % de los proyectos quiere decir que son actividades de mucho menor nivel de recursos, porque constituyen el 49 % de ellos, pero, en cualquier caso, es la mitad de los recursos de la universidad. Es un dato muy importante desde el punto de vista del arranque, pero es un dato muy malo desde el punto de vista de la consolidación. Hay que evaluar que algo igual no está funcionando del todo bien, porque ha funcionado bien para el arranque, pero ahora no funciona tan bien para la consolidación.

Hasta ahora estamos hablando sólo de financiación pública, y la financiación del Gobierno vasco es de este tipo, es decir, del conjunto de financiación de la que se destina una parte a investigación. Pero el propio Gobierno vasco tiene su propio modelo de financiación en paralelo.

Ahora fijémonos en la captación de recursos externos. En ciencias sociales, en términos comparativos, recordemos que estábamos aproximadamente en el volumen de recursos públicos en torno a un 10 % aproximadamente y un 7 % o un 8 % en humanidades. Por lo tanto, en ciencias sociales ello te obliga a captar de recursos externos, es decir, vincularse al mercado. En ciencias sociales, en términos relativos, su posición está a medio camino entre la financiación pública y la vinculación con el mercado. Actualmente estamos prácticamente en el 30 % de los recursos de investigación derivados desde el mercado frente a los públicos.

Estos son los contratos externos en Ciencias Sociales. Solamente los dineros provenientes de contratos externos, no de financiación pública, aproximadamente 400 millones, de los cuales la mitad provienen de la Administración autonómica, el propio Gobierno y la gestión administrativa. En nuestro caso, está la Diputación Foral, con un 14 %, y la empresa privada, con otro 14 %, que sería administración local, empresa pública, que no es expresamente mercado privado, pero es el mercado.

Ahora veremos el cúmulo de proyectos y la inversión en ciencias sociales en el conjunto de los cinco años por áreas de actividad. Lógicamente, Economía se lleva la mayor partida, luego están Psicología y Derecho con una aportación muy similar, a continuación tenemos el área de Política, que también depende mucho del grado de consolidación. Es también interesante ver estos datos en términos de ratio, porque no es lo mismo un departamento o un área de conocimiento con cincuenta profesores que uno que sólo tiene diez. Ahora hablaríamos de la utilidad en captación de recursos por áreas, con lo cual cabría destacar también un área pequeña pero productiva. Porque si a esto le añadimos la juventud de los departamentos, ello da la idea incluso de un cambio generacional y de un cambio de orientación disciplinaria hacia las tareas de investigación o la captación de recursos, sabiendo que todos los que estamos ahí tenemos que compartir una cosa con la otra, y con la educación y la enseñanza que en carreras nuevas requieren una atención especial. Pero, claro, la ironía de lo que aparecía como fuente principal de recursos aparece casi la última de la fila solamente en el ratio de productividad.

Vamos a ver este mismo dato pero en valores exactos. Fijaos que aquí los economistas y Psicología se llevan la gran parte de la tarta en la captación de los contratos externos, que son 400 millones. Los economistas tienen, a la hora de obtener recursos del mercado, más facilidad, probablemente también por los propios *outputs* de su disciplina en proyectos competitivos. Los indicadores son de rendimiento de publicación. Aquí tenemos tanto datos absolutos como porcentajes, o sea la productividad convencional. Porque no estamos hablando de bases de datos, no estamos hablando de Internet, no estamos hablando de otras fuentes de difusión, sino de difusión referente a libros, a capítulos de libros, a artículos de revistas, a ponencias y a congresos, etc., donde vemos también la productividad, lo que esta inversión a cinco años ha generado desde el punto de vista de la rentabilidad en términos académicos. Aquí tenemos los ratios por profesor en cada área de conocimiento, donde se ven de nuevo algunas áreas que destacan en la captación de recursos, y no tanto en la productividad académica, aunque hay una diferencia grande entre los distintos indicadores. No entramos en la calidad, evidentemente no estamos mirando aquí si se trata de revistas de alto impacto o de menor impacto. Y luego hay un tema importante que tiene que ver con la consolidación de los grupos emergentes: nuestra mezcla de evaluación ligada a sexenios, y por lo tanto a complementos salariales. Ahí tenemos la relación entre la categoría profesional de los investigadores principales y los proyectos. Fijémonos, los catedráticos destacan clarísimamente en el liderazgo de los proyectos de investigación, pero no destacan en los liderazgos de los contratos externos. Es decir, o los criterios de evaluación están hechos por catedráticos para catedráticos y son un tapón, o los jóvenes son mucho más dinámicos y más abiertos a las nuevas necesidades del mercado y tienen mucha más vinculación con el mercado que los seniors del sistema, o un poco de todo. Ello obliga a repensar la evaluación y los criterios de investigador principal y otra serie de consideraciones que se deben tener en cuenta a la hora de formarlos o de establecer la formación de los

equipos. Es mucho más probable que tenga que ver con la colaboración de los equipos que con el investigador principal.

Y aquí tenemos la misma dimensión, la de la categoría contractual, entonces, como véis, el modelo es muy similar en todos los campos. Aquí hay uno de los datos de volumen, lo cual es lógico en una universidad donde el volumen de catedráticos es menor que el de titulares. Éstas son las opiniones, los datos básicos, sobre cual es el estado bruto de los recursos de la investigación en Ciencias Sociales, pero vamos a ver algunas opiniones sobre esta cuestión de los recursos, de las evaluaciones, para establecer la diferencia entre unos y otros.

En primer lugar, ya se ha visto que hay un acuerdo bastante generalizado sobre la necesidad de la evaluación permanente de la investigación en los departamentos, pero tenemos que fijarnos en la diferencia entre el grado de adhesión a esa idea entre los catedráticos y los titulares, lo cual vuelve a dar la pista de que los titulares no ven tan claro que lo que se está haciendo se esté haciendo bien, o que el elitismo que se está aplicando sea realmente un elitismo eficaz, y no un elitismo puramente corporativo, de tapón.

¿Qué le parecería que se hiciese realmente una evaluación de la investigación en el departamento? Es decir, que se evalúe la productividad de los catedráticos en sí, porque los titulares probablemente no están en contra de la evaluación, sino que están en contra de las inercias de evaluación que tenemos en este momento. Y es importante ver también cuales son los factores que influyen en la investigación, porque, como vemos, ésta es la diferencia: cuales son los factores que influirían en esta decisión. En primer lugar, la dotación de personal y aquí sí que no hay diferencias entre categorías, aunque los contratados están un poco menos motivados que los catedráticos y titulares. En cualquier caso, vemos que es una protesta lógica, es decir, se plantea el problema de cómo alguien se puede dedicar a la investigación si tiene un contrato precario.

En cualquier caso, la investigación es casi una cuestión de vocación, luego vienen los recursos, obviamente, y viene la motivación. Es un elemento cualitativo que, insisto, en evaluación hemos de tener muy en cuenta, porque los criterios objetivistas desmotivan. ¿Cuál cree usted que debería ser el objetivo de los criterios de evaluación en investigación? Y aquí vemos que nos dividimos entre racionalizar los recursos y facilitar más medios. Una cosa es que lo que tenemos se aplique bien para dar más medios a algunas ideas. Y aquí vemos de nuevo la diferencia entre unos y otros, entre catedráticos y titulares, aquí sí que hay diferencia significativa. Racionalizar los recursos para los catedráticos, que son los que los administran, o los administramos, y facilitar más medios, dicen los titulares, que son los pilares de esto. Ellos están más motivados y los otros son los que administran. Pero esta diferencia existe, y no por otra razón. No podemos pensar que sea un planteamiento puramente corporativo. Y luego hay también el establecimiento de los criterios, que es lo que discutimos, es decir, cuales son los criterios que se deberían de tener en cuenta. Está lo que habíamos discutido del Social Citation Index, pero yo tengo una cierta reticencia a aplicar tales criterios a pesar de que otros están demasiado convencidos de que eso ores matices...

Creo que hay que diferenciar, en primer lugar, el escenario, aparte de la distinción que he dicho de los grupos, considero que no se puede aplicar el mismo criterio de evaluación, sea cual sea el grado de maduración de los grupos, porque entonces estaríamos siempre aplicando criterios de excelencia y los jóvenes tendrían que irse directamente al mercado, porque no van a tener ninguna posibilidad de obtener recursos para sus ideas, que seguro que son buenas, aunque no estén maduras. Por lo tanto, lo que hay que distinguir son los escenarios de la evaluación, es decir, una cosa es la preevaluación, que es el hecho de pararse de vez en cuando a ver lo que estamos haciendo, porque, como os decía anteriormente, los criterios de evaluación acaban condicionando las estrategias de evaluación. Los criterios de evaluación están montados sobre una estrategia de promoción de conocimiento, es la pescadilla que se muerde la cola. Al final, la producción científica es clónica a los criterios que hemos aplicado, sin saber si son adecuados o no. Es decir, si no sabemos qué es lo que estamos haciendo, cómo estamos aplicando unos recursos, no sabemos qué productividad nos están dando y con qué criterios de productividad se aplican que es quizás lo más importante.

Segundo, hay que remarcar que existe mucho cinismo en esto de la evaluación. Yo lo he visto desde dentro. Nos ponemos muy dignos, muy corporativos, pero hay que reconocer que hay mucha falta de criterios, o muchos criterios distintos, y algunos no del todo tan racionales como aparentan. Al final resulta ser malo, porque se compensan los déficits en evaluación. Pero, en fin, uno que lo ve por dentro..., ve que se sale como se puede. Y, además, algo muy importante es que una cosa es la evaluación de las ideas, de las propuestas, y otra cosa es la evaluación de los recursos. Y que, además, la evaluación de las propuestas se hace en compartimentos estancos, y sin embargo la asignación de los recursos no se hace en compartimentos estancos..., sino que en un mismo criterio los estándares no son exactamente los mismos, por ello la racionalidad tiene algunos déficits en su aplicación. Entonces, si no nos paramos de vez en cuando a evaluar qué es lo que hemos hecho, en qué hemos empleado los recursos y el grado de cumplimiento de los objetivos, que yo creo que es un criterio fundamental en evaluación, estaremos en una inercia muy cínica de que, evidentemente, algo se produce, bastante se produce, pero no sabemos si se produce lo que queremos que se produzca o si se produce lo que los investigadores dicen que se pretende producir.

Luego está el tema de la evaluación de las propuestas, que creo que es lo que hemos venido a discutir aquí, es decir, cómo se evalúan las propuestas para que tengan éxito. Unas veces decimos que es la evaluación *ex-post* en la medida en que no estemos en grupos emergentes, sino que estemos valorando ya la consolidación de grupos y sobre todo los grupos de excelencia, donde lo que se evalúa es una propuesta de un plan de investigación. Ahí lo que se evalúan son fundamentalmente los resultados, puesto que ya tiene asegurada la financiación para diez años por ser un grupo de excelencia. Y entonces los criterios no pueden ser los mismos para los grupos emergentes con ideas emergentes, porque en el grupo de excelencia lo fundamental es la coherencia del

grupo, el grado de internación del grupo, la captación de los recursos externos, la plurinomialidad de las publicaciones, porque eso significa que el grupo existe. En los grupos emergentes las publicaciones son individuales, es decir, un grupo de investigación existe o es una suma de individuos que ponen los recursos para elevar el ratio en la asignación de los recursos posteriores. Varía si estamos hablando de un grupo que tiene continuidad, que esta publicando conjuntamente, que aparece en las reuniones, que está en las revistas, etc. y, en definitiva, que tiene un plan de investigación, no tiene una idea, tiene un plan, que es una cosa distinta, donde pueden haber muchas ideas que se van realizando en el tiempo.

Un nevo tema que considero importante es la existencia de líneas estratégicas. En nuestra universidad, un criterio que hemos introducido es el criterio de seguridad de los grupos ya consolidados frente a los grupos emergentes. Así vemos que en los grupos emergentes hay un chaval o una chavala con calidad científica que tienen una cierta trayectoria, que el equipo tiene cierta solvencia, que el plan de trabajo está bien definido, los objetivos están bien definidos, la idea está bien, pero nada dice que aquello sea realizable, no hay seguridad de que se vayan a obtener los objetivos que tienen previstos, pocos o muchos. Sin embargo, en el caso de los grupos más consolidados ésa no es la cuestión, no es la seguridad de los resultados, a veces es la inseguridad de los resultados, es decir, es una idea de alto riesgo que no se sabe si va a conseguir un buen resultado, pero en ese grupo sí nos podemos arriesgar a fracasar, porque están en un nivel más límite en la obtención de resultados, porque ellos harán una aventura más seria. Además, como la aventura no la financiamos nosotros sino que la compartimos, con el mercado o con quien sea, es un riesgo que hay que correr, que en ciencias experimentales es más lógico, porque es precisamente sobre lo que se trabaja, pero en ciencias sociales era hasta ahora la seguridad.

Otro aspecto importante en la valoración de los grupos desde el criterio más cualitativo es la capacidad formativa. Un grupo está más consolidado cuanto más capacidad de formación tiene, es decir, está generando una cartera de investigadores en la que se están haciendo tesis doctorales. Es curioso como cambian los criterios, hasta hace muy poco tiempo el criterio era financiar infraestructura, y ahora es razonar la investigación a ver si sabe lo que quiere.

Otro aspecto importante, a mi modo de ver, acerca de la categoría de las revistas, que a mi me parece que no es lo más importante en la difusión, es cuantos grupos de investigación generan bases de datos disponibles para la comunidad científica. ¿Alguien se ha preocupado de averiguarlo?, porque el conocimiento no es solamente la evaluación que yo haga de mis resultados, sino también que mis datos estén en disponibilidad en un banco de datos para la comunidad científica, porque probablemente ahí está el mayor valor añadido de la información que yo he captado y que he ordenado. El análisis es una parte nada más, el gran esfuerzo inversor probablemente está en la captación de los recursos, sobre todo en ciencias sociales. Los que hacemos investigación más aplicada sabemos que la mayor parte de los recursos que se nos

asignan son fundamentalmente para la captura de información, que queda absolutamente oculta. Administrados celosamente por nosotros mismos, para poder publicar nosotros solos, entonces, el esfuerzo de disponer de bases de datos abiertas a la comunidad científica creo que es un elemento de calidad importantísimo.

Yo tenía preparados unos datos sobre las dimensiones más jerarquizadas de la evaluación de los proyectos, digamos más precompetitivos, pero en el debate surgirán otras ideas que me parecen más interesantes en cuanto a lo que son las expectativas.