



INSTITUT DE FRANCE
Académie des sciences



**L'ÉVALUATION INDIVIDUELLE DES CHERCHEURS
ET DES ENSEIGNANTS-CHERCHEURS
EN SCIENCES EXACTES ET EXPÉRIMENTALES**

**Rapport remis le 08 juillet 2009
à Madame la Ministre de l'Enseignement Supérieur et de la Recherche**

SOMMAIRE

Principales recommandations	p.01
Résumé et recommandations	p.03
Remarques liminaires	p.09
Introduction : un état des lieux critique	p.09
1. Les règles de déontologie et d'éthique	p.11
2. Les critères et outils de l'évaluation de l'activité de recherche	p.13
2.A - Introduction : la place primordiale de l'évaluation qualitative	p.13
2.B - L'évaluation quantitative des travaux de recherche : la bibliométrie	p.13
2.C - Les autres critères d'évaluation de la recherche	p.16
2.D - Les critères et outils de l'évaluation des activités d'application de la recherche	p.17
2.E - L'évaluation par les pairs	p.18
3. Les critères d'évaluation des activités d'enseignement	p.19
3.A - L'enseignant-chercheur et l'équipe pédagogique	p.19
3.B - Les expériences pédagogiques	p.20
3.C - L'évaluation de l'enseignement	p.21
4. L'évaluation des activités d'intérêt collectif	p.22
5. La fréquence et la forme des évaluations	p.23
6. Les Evaluateurs	p.25
6.A - Le choix des évaluateurs	p.25
6.B - La place respective de l'évaluation nationale et de l'évaluation locale	p.26
7. L'utilisation de l'évaluation	p.26
7.A - Les recrutements	p.26
7.B - Les promotions	p.27
7.C - Le suivi de carrières	p.27
7.D - La mise en place d'un dispositif de suivi des carrières	p.28
Conclusion	p.28
Annexe 1 : Composition du groupe de travail	p.30
Annexe 2 : Liste des personnalités auditionnées	p.32
Annexe 3 : Glossaire	p.33
Annexe A : Où trouver l'information de base ?	p.34
Annexe B : Des indicateurs à éviter	p.35
Annexe C : Indicateurs de productivité	p.36
Annexe D : Indicateur d'excellence dans la production	p.38
Annexe E : Biais et perversions	p.39

L'évaluation individuelle des chercheurs et des enseignants-chercheurs en sciences exactes et expérimentales

Principales Recommandations

Pour une déontologie incontestable

- 1 - Assurer la qualité des évaluateurs à partir de pré-requis définis par l'AERES
- 2 - Imposer une déclaration obligatoire des conflits d'intérêt directs et indirects, signée par les évaluateurs, et remplacement au sein du comité si nécessaire
- 3 - Limiter la durée des mandats à trois ans avec rotation annuelle par tiers pour dissocier les lobbies

Évaluation de l'activité de recherche

- 4 - Réaffirmer la valeur primordiale de l'évaluation directe des travaux scientifiques à partir des publications originales et éventuellement d'auditions
- 5 - Encadrer la bibliométrie en s'assurant qu'elle soit utilisée de façon rigoureuse, en évitant tous les biais trop souvent négligés, en adaptant les indicateurs selon les disciplines, et en l'intégrant à l'évaluation qualitative. Encourager le développement d'une recherche critique de nouveaux indicateurs
- 6 - Donner la place qu'elle mérite à la recherche appliquée pour ceux qui s'y consacrent
- 7 - Prendre en considération les critères objectifs de notoriété scientifique nationale et internationale

Évaluation de l'activité d'enseignement

- 8 - Etablir une grille d'évaluation de l'enseignement en donnant une large place à l'appréciation des étudiants
- 9 - Intégrer des critères objectifs tels que le contenu et l'originalité de l'enseignement ou la production d'ouvrages pédagogiques

Évaluation des charges d'intérêt collectif

- 10 - Prendre en compte les charges d'intérêt collectif, plus particulièrement celles impliquant des responsabilités importantes

Autres mesures importantes

- 11 - Réduire le nombre des évaluations au strict nécessaire, en distinguant les évaluations approfondies (recrutements et autres étapes-clés de la carrière), des évaluations de suivi

- 12 - Procéder aux recrutements en deux étapes : admissibilité sur dossier, admission après audition et examen approfondis
- 13 - Toute évaluation approfondie ne doit être entreprise que si elle peut conduire à des conséquences significatives pour l'intéressé. Créer des structures de ressources humaines en vue d'une aide à une évolution de la carrière des personnels en difficulté
- 14 - Bien mettre en perspective le rôle de l'évaluation nationale (essentielle pour l'activité de recherche) et de l'évaluation locale (la mieux placée pour l'évaluation de l'enseignement et celle des charges d'intérêt collectif) en veillant néanmoins à ce que les structures locales d'évaluation incluent un pourcentage suffisant d'experts extérieurs.

L'évaluation individuelle des chercheurs et des enseignants-chercheurs en sciences exactes et expérimentales

Résumé et Recommandations

Déjà pratiquée depuis longtemps au niveau des unités de recherche et des chercheurs, l'évaluation est désormais considérée comme naturelle par l'ensemble de la communauté scientifique et sa généralisation est devenue d'actualité, notamment en raison de la loi LRU qui confère un pouvoir nouveau aux Établissements à l'échelle locale. Elle s'inscrit dans le difficile contexte universitaire lié à l'hétérogénéité des conditions de travail des enseignants-chercheurs qui concerne aussi bien les infrastructures et les équipements universitaires que le niveau des étudiants et des filières de formation. Dans ce contexte particulier, l'Académie a jugé important de formuler des recommandations qui s'appuient sur trois grands principes : compétence, transparence et éthique.

1. Les règles de déontologie et d'éthique

- La durée du mandat des évaluateurs devrait être courte (3 ans) avec renouvellement des membres des comités par tiers chaque année.
- Un regard extérieur à la discipline est souhaitable ainsi que le recours à des experts extérieurs à l'établissement et éventuellement étrangers, dans des proportions très significatives (la loi LRU précise 50% pour les recrutements).
- La procédure et les critères utilisés pour l'évaluation individuelle doivent être publiés, avec des adaptations propres à chaque discipline ou sous-discipline, tant au niveau national qu'au niveau de chaque Établissement ou organisme.
- Un effort particulier doit être fait pour identifier les conflits ou les communautés d'intérêt qui ne sont pas immédiatement apparents, et tout problème de déontologie doit être déclaré à l'avance.
- Chaque membre d'un comité d'évaluation doit conserver une stricte obligation de confidentialité, seul son président étant habilité à donner des indications plus précises en cas de contestation.
- Les rapports doivent être transmis dans leur intégralité aux intéressés, sans modifications, sachant que la confidentialité vis-à-vis de l'auteur sera préservée par le Comité d'évaluation et son président.
- Un véritable engagement d'éthique personnelle doit être signé par chaque évaluateur. Tout manquement aux règles de déontologie sera considéré comme une faute professionnelle grave.

2. Les critères et outils de l'évaluation de l'activité de recherche

L'évaluation de la recherche doit intégrer plusieurs niveaux d'analyse.

Évaluation qualitative

Fondée sur l'analyse directe des travaux scientifiques, éventuellement complétée par une audition réalisée dans des conditions de temps adéquates, l'analyse qualitative est l'élément essentiel de l'évaluation approfondie. Les indicateurs bibliométriques et autres critères quantitatifs ne peuvent pas se substituer à cette évaluation assurée par les pairs mais peuvent, une fois bien développés et compris, devenir un élément d'aide à la prise de décisions.

Évaluation quantitative

Les indicateurs bibliométriques peuvent se révéler très utiles lorsqu'ils sont bien utilisés, replacés dans le contexte de la discipline, et intégrés à l'évaluation qualitative.

- Les indicateurs bibliométriques ne peuvent pas être utilisés isolément pour effectuer un classement.
- Il faut donner plus d'importance aux citations des articles qu'au facteur d'impact des journaux où ils sont publiés (sauf pour le cas des jeunes chercheurs). Les index H et G, fondés sur les citations, sont utiles mais ont leurs limites et demandent à être complétés par de nouveaux indicateurs.
- Il faut tenir compte, dans chaque citation, du nombre d'auteurs et, dans les disciplines où l'ordre des auteurs n'est pas alphabétique, de la place de l'auteur considéré.
- L'Académie propose d'engager une réflexion inter-organismes et inter-disciplinaire, en liaison avec l'OST et l'AERES, qui devra se pencher sur l'utilisation des outils bibliométriques et la création de nouveaux indicateurs. Il faudrait valider les outils actuellement utilisés, par des tests rétrospectifs.

Autres critères de notoriété

La qualité scientifique d'un chercheur peut être évaluée sur de nombreux autres critères que ceux fondés sur la bibliométrie, en particulier la capacité à organiser, coordonner, diriger des recherches, la rédaction de livres, la rédaction d'ouvrages de vulgarisation, le nombre de langues dans lesquelles ils sont traduits, les invitations en tant que conférencier en séances plénières, la direction de programmes internationaux ou la présidence d'une union scientifique internationale, la participation à des comités ou agences internationales, les postes de rédacteur en chef de revues internationales, l'attribution de contrats importants, la reconnaissance par des prix ou distinctions nationales ou internationales, les élections à des académies françaises ou étrangères et d'autres distinctions telles qu'une nomination à l'IUF,

l'organisation d'écoles d'été, de symposiums et de congrès internationaux de haut niveau, etc..

Évaluation des applications de la recherche

Contrairement à ce que l'on constate à l'étranger, et bien que des progrès aient été faits dans ce domaine, les réalisations et applications industrielles sont insuffisamment prises en compte dans l'évaluation des chercheurs en France.

- La valorisation doit devenir un critère essentiel d'évaluation pour ceux qui s'y consacrent et être un facteur de promotion au même titre que les publications.
- Il serait important de construire la grille d'évaluation en mettant en bonne place la question de la pertinence de la recherche.
- Il faudrait aussi définir des critères d'évaluation pour les réalisations de la recherche qui ne comportent pas d'applications immédiates directes comme par exemple les logiciels ou les prototypes, mais qui n'en sont pas moins importantes.

In fine, il revient aux comités de pairs de faire la synthèse des éléments d'évaluation qui viennent d'être mentionnés, en privilégiant l'analyse directe des travaux à partir des articles originaux ou d'auditions. Cette évaluation doit bien sûr intégrer les indicateurs quantitatifs mais aussi pleinement prendre en compte l'originalité de la recherche et sa pertinence.

3. Les critères d'évaluation des activités d'enseignement

La loi LRU et le récent décret du 23 avril 2009 fixant les dispositions statutaires applicables aux enseignants-chercheurs instaurent l'obligation d'évaluer leurs trois types d'activités : recherche, enseignement et activités d'intérêt collectif. Ceci s'avère d'autant plus nécessaire que l'importance relative de ces trois activités peut varier au cours de la carrière. Concernant l'évaluation de l'enseignement sous toutes ses formes, l'Académie préconise les recommandations suivantes :

- L'évaluation des activités d'enseignement peut être réalisée selon plusieurs approches conduisant à la production d'une grille d'évaluation aux niveaux national et local, en sachant que le niveau local est évidemment le plus pertinent. Un critère important est l'appréciation des étudiants sur les enseignements, en tenant compte du fait qu'elle peut exposer à des effets pervers.
- L'évaluation des activités d'enseignement doit aussi intégrer certains critères
- objectifs comme le contenu et l'originalité de l'enseignement (publications à caractère pédagogique, ouvrages édités, mise en ligne de cours ou TD-TP, expositions, etc.).
- Un relevé des charges annuelles d'enseignement doit être publié chaque année par chaque établissement, les éventuelles décharges d'enseignement devant être clairement mentionnées et justifiées.

- L'établissement doit publier, chaque année, le pourcentage de succès et le devenir des étudiants par niveau d'étude (L1, L2, L3, M1, M2, doctorat), les débouchés des étudiants par filière et à l'issue de la thèse. Ces éléments devront être pris en compte, dans la mesure du possible, pour l'évaluation individuelle des enseignants-chercheurs.
- Les meilleurs Maîtres de Conférences qui s'investissent très fortement (en formation initiale et en formation continue) et sont unanimement reconnus pour leurs qualités pédagogiques, devraient pouvoir être candidats à des promotions locales comme la Hors Classe, ou bénéficier d'une prime exceptionnelle à l'échelle de chaque établissement. A l'inverse, ceux qui négligent leurs tâches d'enseignement devraient en subir des conséquences.

4. L'évaluation des charges d'intérêt collectif

- Les charges administratives et d'intérêt collectif doivent être prises en compte dans l'évaluation des enseignants-chercheurs et des chercheurs, conformément aux dispositions des nouveaux textes en vigueur, notamment celles qui impliquent de véritables responsabilités (coordonnateur de 1^{ère} année L1, responsable d'UFR, missions de coopération internationale, orientation et insertion professionnelle des étudiants, coopération avec les entreprises, valorisation, diffusion de la culture scientifique et technique, etc.).
- Un relevé des charges annexes devrait être publié par l'établissement pour chaque enseignant-chercheur.
- Cependant, dans l'état actuel, il n'existe pas de critères objectifs pour les évaluer. Il faudrait donc élaborer une grille de critères d'évaluation à cet effet.

5. La fréquence et la forme des évaluations

La fréquence actuelle des évaluations est trop grande. L'Académie fait plusieurs recommandations :

- Il faut distinguer les évaluations « approfondies » des évaluations de suivi.
- Le nombre et la fréquence des évaluations approfondies doivent être réduits aux seules étapes importantes de la carrière d'un chercheur ou d'un enseignant-chercheur, à savoir les recrutements et les changements importants de grade ou de corps.
- Le recrutement est une étape clé du dispositif, compte tenu du statut de fonctionnaire d'Etat des personnels recrutés.

*Il faut généraliser un système à deux tours avec une admissibilité sur dossiers et une admission à l'issue d'une présentation orale significative et d'un entretien approfondi. Les critères de créativité et d'originalité des recherches requis pour un recrutement étant difficiles à apprécier par le simple jeu d'une analyse bibliométrique ou d'une audition préparée à l'avance, l'examen approfondi par les pairs doit jouer pleinement ce rôle.

L'aptitude à l'enseignement doit également être testée à l'aide de séminaires adaptés.

- Les listes de qualification nationales créées pour pallier l'hétérogénéité de niveau des thèses et des HDR ne sont pas pleinement satisfaisantes. L'autonomie des établissements doit conduire à une importante redéfinition des critères nécessaires à leur obtention et aux rôles essentiels que doivent jouer les écoles doctorales d'une part, les conseils scientifiques d'Établissement de l'autre. Ces recommandations doivent aboutir rapidement à la délivrance de grades reconnus par leur qualité.

*Quatre étapes d'évaluation, liées aux inscriptions préalables sur les listes de qualification, sont nécessaires au passage comme professeur. Ceci est à comparer aux deux étapes requises dans les pays comparables au nôtre. Une réflexion globale à ce sujet doit être réalisée.

- Limiter les évaluations de suivi au Contrat quadriennal d'Établissement, sur la base d'une fiche simplifiée destinée au suivi normal de l'activité des personnels

6. Les évaluateurs

- La qualité scientifique des évaluateurs est fondamentale. L'Académie propose pour chaque catégorie d'évaluateurs, la publication de pré-requis qui pourraient être préparés par l'AERES.
- Des mesures importantes doivent être prises pour assurer la motivation des meilleurs évaluateurs en leur facilitant la tâche, en réduisant la durée des mandats et en prévoyant la prise en compte de cette activité dans les charges administratives ou d'intérêt collectif
- Tous les évaluateurs devront être évalués pour garantir leur compétence.
- Les comités d'évaluation propres à chaque université doivent être gérés localement mais inclure une fraction significative de membres extérieurs (comme le demande la loi LRU pour le recrutement).
- Bien mettre en perspective le rôle de l'évaluation nationale (essentielle pour l'activité de recherche) et de l'évaluation locale (la mieux placée pour l'évaluation de l'enseignement et des charges d'intérêt collectif).

7. L'utilisation de l'évaluation

- Une des difficultés majeures du système d'évaluation est l'absence fréquente de conséquences de l'évaluation. Il faut donc veiller à ce qu'elle ne soit réalisée que lorsqu'elle peut conduire à une promotion ou à une réorientation.
- Il faut dissocier l'évaluation de la qualité d'une activité, de la gestion des carrières et ne pas mélanger évaluation et ré-orientation.
- Un suivi de la carrière devrait être fait grâce à un Comité ad-hoc qui devrait jouer le rôle d'un coordinateur de carrière. Ce dispositif de bonne gestion des ressources humaines devrait être mis en place à l'échelle de chaque établissement.

- Les Établissements devraient afficher leurs orientations scientifiques et pédagogiques de manière à ce que les appels à candidatures ne soient pas biaisés.
-

L'évaluation individuelle des chercheurs et des enseignants-chercheurs en sciences exactes et expérimentales

Remarques liminaires

Comme cela est indiqué dans le titre, ce rapport ne concerne que les disciplines scientifiques exactes ou expérimentales, des mathématiques à la biologie et aux sciences médicales et pharmaceutiques. Le rapport ne prend pas en considération les sciences humaines et sociales qui, certes, partagent avec les autres sciences nombre de problèmes d'évaluation, mais ont aussi sur plusieurs points leurs propres spécificités.

Le rapport concerne à la fois l'évaluation des chercheurs et des enseignants-chercheurs. Il doit être clairement compris que lorsqu'il sera référé aux *Établissements*, il s'agira des universités et lorsqu'il sera question d'*organismes*, il s'agira des organismes de recherche. Le rapport concerne aussi les Établissements ou organismes dépendant d'autres ministères que celui de l'Enseignement Supérieur et de la Recherche.

Enfin, il faut rappeler que l'évaluation collective est hors du champ de ce rapport, mais il ne faut pas sous-estimer l'intrication de l'activité de chaque chercheur ou enseignant-chercheur avec celle du groupe auquel il appartient, laboratoire ou équipe.

Introduction : un état des lieux critique

L'évaluation individuelle est un acte essentiel qui permet d'une part d'évaluer les compétences des chercheurs et enseignants-chercheurs et les résultats de leurs recherches, et d'autre part d'asseoir les recrutements, les promotions et le suivi de l'activité tout au long de leur carrière, sur des bases objectives. Or, contrairement à ce qui se passe dans d'autres grands pays, il n'y a pas en France de véritable consensus sur l'évaluation, même si celle-ci - et cela est un progrès important - est aujourd'hui pleinement acceptée par l'ensemble de la communauté. Ce problème a acquis une actualité particulière en raison de la loi LRU qui confère un pouvoir nouveau aux établissements à l'échelle locale. Plusieurs aspects doivent être considérés.

Tout d'abord, pour qu'une évaluation soit acceptée, elle doit s'appuyer sur les trois principes fondateurs de *compétence*, de *transparence* et d'*éthique*, et être effectuée sous le contrôle des pairs.

La *compétence* des évaluateurs est un pré-requis indispensable à toute bonne évaluation et doit être garantie aussi bien au niveau local que national, ce qui n'est pas toujours le cas. De même, la *transparence* des évaluations, notamment pour ce qui concerne le rapport des évaluateurs, n'est pas toujours assurée. Enfin, les modalités d'évaluation, en particulier les règles de bonnes pratiques et d'*éthique*, ne sont pas la plupart du temps clairement affichées. Cette insuffisance de la déontologie peut conduire à des conflits ou à des communautés

d'intérêt d'individus ou de groupes, donnant trop souvent une mauvaise image de l'évaluation.

Il y a paradoxalement plutôt trop d'évaluations que pas assez dans le système français actuel : recrutement, listes de qualification, changements de corps, promotions au cours de la carrière, demandes de financement, etc.. Cette trop grande fréquence fait perdre beaucoup de temps et d'énergie, et nuit à l'impact de l'évaluation elle-même.

L'introduction de critères quantitatifs dans l'évaluation de l'activité scientifique par la bibliométrie représente un progrès. Néanmoins, elle peut donner lieu à des biais importants liés à la fois à la mauvaise utilisation des indicateurs et à la place excessive qui lui est donnée par rapport aux autres critères d'évaluation. Il est donc nécessaire de lever la confusion qui existe souvent entre évaluation et bibliométrie, et rappeler que tous ces critères n'ont de sens qu'à travers une évaluation finale par les pairs.

A l'exception des recrutements, des changements de corps et des promotions, les évaluations individuelles sont trop rarement suivies d'effet, ce qui conduit à une certaine lassitude des experts sollicités et à de fréquents refus de participer de ces derniers, notamment de la part de ceux qui sont les plus occupés et sont souvent les plus compétents.

A cela s'ajoute un point qu'il faut souligner et qui sera récurrent dans ce rapport, l'évaluation se présente dans un contexte très différent selon les disciplines.

S'agissant des activités d'enseignement et des actions d'intérêt collectif, la situation est encore plus complexe en raison de l'insuffisance de critères objectifs d'évaluation.

Toutes ces difficultés sont aggravées par des procédures administratives lourdes, et coûteuses, y compris en temps passé par les évaluateurs et les évalués, pour une efficacité limitée et des résultats souvent critiquables.

Enfin, enseignants-chercheurs et chercheurs des établissements d'enseignement supérieur et de recherche publics ont un statut de fonctionnaire d'État, ce qui complique le suivi de l'évaluation individuelle, notamment pour l'octroi de promotions, de primes particulières ou encore de décharges de service d'enseignement, soit à l'échelle locale (Établissements essentiellement), soit à l'échelle nationale (CNU ou organismes, Commission Nationale des Enseignants-Chercheurs dépendant du ministère de l'Agriculture (CNECA), ou autres comités nationaux).

Par ailleurs, l'ensemble de cette situation s'inscrit dans le difficile contexte universitaire lié à l'hétérogénéité des conditions de travail des enseignants-chercheurs concernant aussi bien les infrastructures et les équipements universitaires que le niveau des étudiants et des filières de formation, conditions d'autant plus délicates que les charges d'enseignement sont lourdes et

qu'un grand nombre d'enseignants-chercheurs n'appartient pas à des formations de recherche labellisées.

Compte tenu de l'ensemble de ces éléments, et dans le cadre du dispositif législatif actuel, avec en particulier la publication de la loi LRU et du nouveau décret du 23 avril 2009 fixant les dispositions statutaires applicables aux enseignants-chercheurs, l'Académie a jugé important de formuler des recommandations, en s'inspirant des meilleures pratiques internationales en la matière et en espérant contribuer ainsi à la création d'un espace européen de l'évaluation.

Ces recommandations traiteront successivement de l'évaluation individuelle sous différents aspects : la déontologie et l'éthique, la recherche, l'enseignement, les activités d'intérêt collectif, la fréquence et la forme des évaluations, et enfin leur suivi.

1. Les règles de déontologie et d'éthique

Une des difficultés majeures de l'évaluation est de savoir éviter, autant que faire se peut, les conflits d'intérêt qui peuvent être subtils car très indirects. Tous les grands pays comparables à la France ont défini des règles de procédure et d'éthique pour pallier ces perversions potentielles. La France est très en retard sur ce point essentiel à la crédibilité et à l'acceptabilité des procédures d'évaluation par la communauté scientifique. L'AERES a mis en place des critères de déontologie et d'éthique mais les objectifs de cette agence d'évaluation concernent les organismes de recherche, les établissements, les laboratoires et les équipes, et non les individus. Rien n'a été élaboré pour l'évaluation individuelle.

S'agissant à proprement parler de la déontologie et de l'éthique, plusieurs points importants doivent être mentionnés :

- La durée d'un mandat d'évaluateur doit être courte (3 ans au maximum) de manière à ce qu'il ne soit pas pénalisé par les règles d'éthique qui lui imposent de ne pas être juge et partie. Ainsi durant son mandat, l'évaluateur ne doit pas être candidat à une promotion, participer au recrutement d'un nouvel entrant ou à une promotion d'un enseignant-chercheur ou d'un chercheur de son équipe. Il ne suffit pas, pour éviter les conflits d'intérêt, que la personne concernée sorte de la salle lorsque l'on traite de son cas car on sait bien qu'elle participera de fait à la décision finale. Une raison supplémentaire pour limiter la durée de ces mandats est que l'on demande aux évaluateurs une étude approfondie des dossiers, ce qui est difficile à imposer sur une longue durée. L'idée de renouveler les comités d'évaluateurs par tiers tous les ans sur trois ans permettrait tout à la fois de conserver un effet mémoire des évaluations et de dissocier les lobbies.

- Un regard extérieur à la discipline est également souhaitable (ex : un physicien dans une

section de biologie, etc.). Il est aussi indispensable d'avoir recours à des experts extérieurs à l'établissement en proportion significative, notamment 50% pour les recrutements comme le stipule l'art. 25 de la Loi LRU. Ces experts extérieurs doivent inclure un certain pourcentage d'étrangers à définir selon les cas, en sachant que ces derniers peuvent être consultés à distance ou par écrit. L'avis des experts étrangers est d'autant plus utile qu'il est fondé sur un jugement comparatif entre plusieurs candidats. Par ailleurs, des lettres d'évaluation personnelles demandées par le Comité sur proposition des candidats sont utiles. Ces modalités doivent s'appliquer à tous les Établissements, quelle qu'en soit la taille, pour éviter à la fois des auto-recrutements ou des promotions locales injustifiées. Les dispositions qui viennent d'être mentionnées concernent les universités mais pourraient être étendues dans leurs principes aux organismes de recherche.

- Pour assurer la transparence, la procédure et les critères utilisés pour l'évaluation individuelle doivent être publiés, avec des adaptations propres à chaque discipline ou sous-discipline, tant au niveau national (sections du CNU ou du Comité national de la recherche scientifique, p. ex.) qu'au niveau local (Établissements essentiellement).

- Un effort particulier doit être fait pour identifier les conflits ou les communautés d'intérêt qui ne sont pas immédiatement apparents. Ainsi, on peut avoir déposé une demande à l'ANR (ex : programme ANR blanc) et être sollicité la même année pour évaluer des demandes dans ce même programme : c'est typiquement un conflit d'intérêt puisqu'il y a compétition dans un même cadre. Si un évaluateur doit rapporter sur un dossier qui lui pose un problème de déontologie ou d'éthique particulier, il doit le déclarer au président du jury en préalable à toute évaluation. Ce dernier décidera alors de la procédure à suivre et du remplacement éventuel de l'évaluateur.

- Chaque membre d'un Comité d'évaluation doit conserver une stricte obligation de confidentialité sur le contenu des débats. Seul son président est habilité, si nécessaire et en garantissant l'anonymat des membres du Comité ayant participé à l'évaluation, à donner des indications plus précises en cas de contestation éventuelle par un personnel évalué. Cela permettra aux évaluateurs de s'exprimer plus librement dans le rapport puisque la confidentialité de l'auteur sera préservée et que l'avis émis sera celui du Comité dans son ensemble. Cette procédure permettrait en outre d'assurer une plus grande transparence. L'Académie estime qu'il est important que les rapports soient transmis dans leur intégralité aux intéressés, sans modifications (sauf pour les remarques ayant une connotation trop personnelle). Elle est en revanche défavorable à la mise de ces rapports sur un site largement accessible.

- Enfin, un véritable engagement d'éthique personnelle doit être signé par chaque évaluateur, comme cela se pratique couramment dans les autres pays. Des règles d'éthique ont déjà été élaborées et peuvent servir de modèle (notamment à l'ERC). Certaines sont également évoquées dans les décrets récents concernant le CNU, sans toutefois être suffisamment

complètes ni faire référence à l'obligation de signature largement pratiquée à l'étranger ; il n'y est pas fait non plus mention de ce qui se passe en cas de manquement. Or, tout manquement aux règles de déontologie et d'éthique doit être considéré comme une faute professionnelle grave.

2. Les critères et outils de l'évaluation de l'activité de recherche

2. A - Introduction : la place primordiale de l'évaluation qualitative

Des évolutions importantes ont récemment marqué l'évaluation de l'activité scientifique : on peut maintenant utiliser des critères quantitatifs fondés sur la bibliométrie c'est-à-dire l'utilisation des publications scientifiques comme indicateurs de la production de leurs auteurs ; l'importance des applications des recherches est plus clairement reconnue ; on a appris à mieux identifier les signes extérieurs de notoriété scientifique à partir des marques de reconnaissance du chercheur par la communauté internationale. Néanmoins, l'évaluation qualitative fondée sur l'analyse directe des travaux de recherche, que certains avaient cru à tort pouvoir être remplacée partiellement par la bibliométrie, reste l'élément essentiel de l'évaluation. Une large place sera pourtant donnée ici à l'évaluation quantitative car il convient de dissiper l'ambiguïté qui existe fréquemment entre bibliométrie et évaluation.

Les différents points qui vont être examinés successivement concernent aussi bien les chercheurs statutaires que les enseignants-chercheurs¹.

2. B - L'évaluation quantitative des travaux de recherche : la bibliométrie

Comme suite au développement de la première grande base bibliographique *Science Citation Index* créée dans les années 1960 (Annexe A), Eugène Garfield a eu l'idée d'utiliser les citations des articles comme critère de délimitation des domaines scientifiques². Ultérieurement, l'utilisation des citations a donné naissance à des instruments de mesure de la visibilité des journaux scientifiques (facteurs d'impact). Plus récemment, l'usage des citations a été étendu à l'évaluation scientifique individuelle grâce notamment à l'utilisation de la base *Web of Science* commercialisée par *Thomson Reuters*, ou *Scopus* commercialisée

¹ L'AERES a déjà proposé une définition qui prend en compte la notion d'enseignant-chercheur publiant en fonction du nombre minimum de publications, ce qui représente une mesure intéressante mais sans doute pas assez discriminante car le seuil proposé par l'AERES pour définir les publiants est trop bas. Toutefois cette valeur est relative et doit être appréciée en fonction des disciplines. Des chercheurs très remarquables peuvent ne pas ou peu publier pendant une période avant de communiquer un résultat majeur, notamment en mathématiques. Enfin, certains chercheurs doivent retarder leurs publications pour ne pas gêner la prise d'un brevet.

² E. Garfield, Citations Indexes for Sciences : A new dimension in documentation through association of ideas, *Science*, 122, 108-111, 1955

par *Elsevier*. La bibliométrie est devenue une pratique qui se développe rapidement. Il est cependant bon de prendre conscience que l'outil bibliométrique reste pour l'instant dans un état rudimentaire.

Lorsque l'on évoque les mesures quantitatives utilisées pour évaluer les travaux scientifiques et leurs auteurs, il y a lieu de bien faire la différence entre l'évaluation des sources que sont les journaux et l'évaluation des auteurs.

- La première est basée sur la mesure de l'impact des journaux identifiée au facteur d'impact *IF* (en anglais) publié par la base *Journal Citation Report (JCR)*, un sous-ensemble de la base *ISI Web of Knowledge* (voir Annexe B). Cette pratique est dangereuse car les revues les plus prestigieuses contiennent aussi des articles de qualité moyenne d'ailleurs peu cités. Un article ne devient pas important par le simple fait qu'il a été publié dans une revue prestigieuse. Toutefois, la critique du facteur d'impact des revues en tant qu'indicateur pour l'évaluation individuelle des chercheurs et enseignants-chercheurs ne doit pas décourager ces derniers de tenter de publier des résultats qu'ils considèrent comme très novateurs dans les journaux de prestige. Il est en effet important que la France soit la plus présente possible dans ces grandes revues qui restent dans une large mesure, en dépit des réserves émises plus haut, une vitrine de la science internationale dans nombre de disciplines.
- La seconde est basée sur la mesure des citations des articles. Elle s'est développée très rapidement depuis 2005 après la proposition par J.E. Hirsch d'un indicateur *h* attribué individuellement à chaque scientifique. Cet indicateur, fondé sur le nombre d'articles ayant un fort nombre de citations, a le mérite de pouvoir être calculé extrêmement rapidement avec le logiciel *Web of Science* de la base Thomson ISI (Annexe C). De plus, un autre indicateur un peu plus élaboré et privilégiant l'excellence, l'indicateur *g*, a été proposé en 2006 par L. Eghhe (Annexe D).

Que penser de l'utilisation de la mesure des citations pour l'évaluation individuelle de la recherche ? Les indicateurs bibliométriques peuvent se révéler très utiles lorsqu'ils sont bien utilisés, c'est-à-dire, replacés dans le contexte de la discipline, prenant en compte la taille de la communauté scientifique qu'elle représente, et intégrés à l'évaluation qualitative (*cf. infra*). La bibliométrie peut donc représenter un outil pertinent si elle est utilisée de manière raisonnée et entre les mains d'experts scientifiques. En tout état de cause, compte tenu de sa large utilisation actuelle et future, il convient moins de la condamner, ce qui serait d'ailleurs excessif, que de souligner l'importance de ses limites et d'améliorer son emploi et sa transparence, ainsi que de proposer de nouveaux indicateurs.

Les indicateurs bibliométriques ne doivent pas être utilisés isolément pour effectuer un classement, à la fois parce qu'ils présentent des faiblesses intrinsèques liées notamment à l'hétérogénéité de la situation individuelle des chercheurs, et à cause de l'importance de l'évaluation qualitative évoquée plus bas. De façon plus analytique, la bibliométrie actuelle

souffre de plusieurs faiblesses :

- La bibliométrie se présente de façon radicalement différente selon les disciplines. Elle est peu utilisée en mathématiques mais démesurément en biologie. En outre, la valeur moyenne des citations et donc des facteurs d'impact et des indicateurs mentionnés plus haut, n'est pas la même dans chaque discipline. Les critères sont donc à définir et à publier pour chacune d'entre elles. Pour une optimisation des données bibliométriques d'un chercheur, il conviendra de ne jamais lui associer un nombre mais de le replacer dans la distribution du domaine homogène auquel il appartient.
- Les indicateurs peuvent engendrer biais et perversions dans le comportement des chercheurs qui pour certains vont plus orienter leur activité vers l'acceptation d'articles dans les revues à fort facteur d'impact que vers une recherche originale et créative (voir Annexe E).
- Selon les disciplines, le nombre d'auteurs varie beaucoup et surtout le choix de l'ordre dans lequel ils figurent est très différent (ordre alphabétique, séniorité, ou autre). Ce point peut poser de sérieux problèmes. Il faut donc tenir compte dans chaque citation, du nombre d'auteurs et, dans les disciplines où l'ordre des auteurs n'est pas alphabétique, de la place de l'auteur considéré.
- Des erreurs sont souvent commises dans le calcul des indicateurs, ce qui est particulièrement grave quand cela est fait à l'insu de l'intéressé. On pourrait proposer qu'une structure telle que l'Observatoire des Sciences et des Techniques (OST) valide les indicateurs calculés.
- Enfin, les indicateurs bibliométriques doivent être appliqués avec beaucoup de précaution aux jeunes chercheurs. Il importe de ne pas utiliser cet outil dans le cas d'activité professionnelle inférieure à 5 ou 10 ans suivant les disciplines.

En résumé, les indicateurs ne peuvent pas se substituer à l'évaluation par les pairs mais peuvent, une fois bien développés et compris, devenir un élément d'aide à la prise de décisions. La constatation d'anomalies (chercheurs de grande valeur ayant publié des articles de grande qualité mais en petit nombre, ou au contraire, chercheurs dont le nombre élevé de citations n'est pas corrélé à la qualité de leur contribution) ne remet pas en question l'ensemble de la bibliométrie mais attire l'attention sur les cas particuliers révélés par les indicateurs bibliométriques justifiant une analyse encore plus approfondie par les comités de pairs. L'Académie est donc favorable à l'emploi raisonné de la bibliométrie dans le processus de l'évaluation individuelle intégré à l'évaluation par les pairs et sous le contrôle des comités d'évaluation. Elle préconise en plus, le développement de bouquets de nouveaux indicateurs afin d'adapter la bibliométrie à la diversité des situations présentées par les différentes disciplines et aussi d'utiliser l'évolution temporelle de ces bouquets d'indicateurs pour suivre l'évolution de carrière des chercheurs. Afin de pallier les faiblesses qui viennent d'être mentionnées, l'Académie des Sciences suggère le lancement d'une réflexion inter organismes

et inter disciplinaire en liaison avec (OST) et l'AERES qui devra se pencher sur l'utilisation des outils bibliométriques et la création de nouveaux indicateurs. Il sera aussi important de valider les outils développés par des tests rétrospectifs.

2. C - Les autres critères d'évaluation de la recherche

La qualité d'un chercheur peut être évaluée sur de nombreux autres critères que ceux fondés sur la bibliométrie comme en témoigne le remarquable formulaire du Conseil Européen de la Recherche (ERC) pour les subventions individuelles, et comme l'avait déjà recommandé l'Académie dans son rapport sur *l'Attractivité des carrières de recherche*. Le nouveau décret du 23 avril 2009 fixant les dispositions statutaires applicables aux enseignants-chercheurs stipule désormais que l'ensemble des activités des enseignants-chercheurs devront être prises en compte. Pour ne s'en tenir qu'à l'évaluation de la recherche, on peut retenir les principaux critères suivants :

- capacité à organiser, coordonner, diriger des recherches (nombre et qualité des thèses dirigées, qualité et devenir des étudiants post-doctoraux, ...)
- rédaction de livres
- rédaction d'ouvrages de vulgarisation,
- écriture de chapitres dans des ouvrages collectifs,
- traduction de ces livres et de ces ouvrages à l'étranger,
- direction de programmes internationaux ou présidence d'une union scientifique internationale,
- cours dans des écoles d'été,
- invitations en tant que présentateur à des conférences en séances plénières,
- postes de rédacteur en chef de revues internationales,
- attribution de contrats, notamment les contrats internationaux,
- reconnaissance par des prix ou distinctions nationales et internationales,
- élections à des Académies françaises ou étrangères, à l'I.U.F., etc.,
- organisation de symposiums ou de congrès internationaux de haut niveau ou de caractère innovant,
- participation à des comités ou agences nationales et internationales,
- actions auprès du grand public.

Le cas particulier de la direction de recherche mérite un commentaire spécifique car elle a plusieurs volets : un volet relatif aux orientations et stratégies de recherche, un volet purement administratif qui peut être lourd, et enfin un volet, souvent délicat, lié au management des ressources humaines (voir aussi § 7 Utilisation de la recherche). Le devenir des étudiants formés au sein du laboratoire de recherche (p. ex. 3 ans après la thèse) doit également être pris en compte dans l'évaluation individuelle au titre de la direction de recherche.

Toutes les fonctions qui contribuent au bon déroulement de la recherche doivent être reconnues par la communauté scientifique. Le travail d'un chercheur lié à l'évaluation ainsi

que ses autres activités d'intérêt général doivent être gratifiants pour la carrière de son auteur. Mais tous ces critères demandent à être interprétés de façon fine et adaptée.

2. D - Les critères et outils de l'évaluation des activités d'application de la recherche

Dans les pays à forte activité économique, les réalisations et applications de la recherche revêtent une importance croissante à cause de leur implication directe dans les produits et processus de haute technologie. La France connaît un certain retard à cet égard, retard que peut contribuer à combler une évaluation des enseignants-chercheurs et des chercheurs qui prenne mieux en compte leur activité dans ce domaine, notamment pour leur promotion.

1° Des attitudes contrastées en matière d'application de la recherche

En France, la recherche fondamentale bénéficie d'un préjugé plus favorable que ses applications, contrairement à ce que l'on constate à l'étranger. La création de prototypes ou de logiciels élaborés est peu valorisée et l'implication d'un chercheur dans les applications n'est pas suffisamment considérée comme un prolongement naturel de la recherche. Ceci produit un décalage entre les connaissances fondamentales et les applications qu'on pourrait en faire. Quand des applications sont entreprises, elles peuvent ne pas aboutir faute d'aide logistique solide, de partenaires industriels engagés, de financements adéquats, et de persévérance. Si la bibliométrie devenait un critère d'évaluation dominant, les chercheurs dont les travaux devraient donner lieu à une valorisation seraient tentés de donner la priorité aux publications en prenant le risque d'empêcher une prise de brevet.

Certes, la situation évolue et des progrès importants ont été faits, grâce à la loi sur l'innovation, la création d'agences de moyens, et la facilité d'obtenir des disponibilités ou détachements de chercheurs pour créer des entreprises. Mais les progrès réalisés restent insuffisants. Les raisons invoquées sont la complexité persistante des procédures, la multiplicité des intervenants et la réticence des institutions de recherche à établir avec l'industrie des partenariats adaptés à la logique économique des entreprises et des investisseurs privés. Une autre raison est la faible prise en compte des réalisations et applications industrielles ou sociétales dans l'évaluation de la carrière des chercheurs. Faire des applications de la recherche, de l'innovation et des autres réalisations scientifiques un critère essentiel d'évaluation des chercheurs qui s'y consacrent est donc une nécessité absolue.

2° Les critères d'évaluation

Dans le domaine *des applications de la recherche*, la question principale est celle de leur utilité scientifique ou industrielle. Chaque chercheur, pour la préciser, pourrait s'aider d'une grille où figureraient certains des critères suivants :

- Dépôt de brevets. Bien que « tout ne se brevète pas », le nombre de dépôt de brevets est un critère essentiel d'évaluation. Il est en fait insuffisant : il faut prendre en compte également les licences d'exploitation et plus encore les résultats, à terme, de l'exploitation de ces brevets.
- Partenariats avec l'industrie et transfert industriel de prototypes.
- Activités de conseil auprès des entreprises.
- Création de start-up, sociétés ou groupements d'intérêt économique.
- Partenariats inter-institutionnels publics ou privés avec les EPST, l'industrie, la santé, etc.
- Implication personnelle du chercheur dans ces applications.

Les réalisations de la recherche, c'est-à-dire toutes contributions scientifiques nouvelles ne comportant pas d'applications industrielles immédiates directes, ne sont pas moins importantes mais sont plus difficiles à évaluer. On se contente généralement d'indicateurs faibles comme par exemple le nombre de téléchargements d'un logiciel libre. L'Académie des sciences préconise de définir des critères d'évaluation de ces activités et de valoriser la contribution des chercheurs qui réalisent des expériences de haut niveau, produisent des logiciels ou des prototypes nouveaux, ou explorent de nouvelles voies d'application sociétale.

L'évaluation est une démarche indispensable pour la carrière des chercheurs et enseignants-chercheurs. Elle peut prendre les formes les plus diverses : promotions, distinctions, primes ou autres intéressements financiers, de façon à introduire dans notre pays une véritable culture d'exploitation de la recherche soulignant le continuum qui doit exister entre la recherche fondamentale et la recherche appliquée, l'une fécondant l'autre et vice versa.

2. E - L'évaluation par les pairs

Comme cela a déjà été dit à plusieurs reprises, on ne peut en aucun cas se limiter à une évaluation quantitative, qu'elle soit fondée sur la bibliométrie, sur les indicateurs de notoriété ou sur les autres critères qui viennent d'être évoqués. Un comité de pairs doit examiner les travaux de recherche en analysant directement leur qualité à partir des articles originaux, ou d'auditions dans la mesure où celles-ci sont réalisées dans des conditions de temps adéquates. Concernant les articles, les évalués devraient choisir un nombre restreint de leurs travaux qu'ils souhaitent soumettre à évaluation (entre cinq et dix). De leur côté, les évaluateurs devraient signer un engagement de lire au moins deux articles parmi ceux proposés par le candidat, comme cela se pratique dans certains pays étrangers comme le Canada. Dans certains cas, surtout pour les experts étrangers, cette évaluation peut être réalisée à distance ou par courrier. Mais il est bien préférable, lorsque cela est possible, qu'elle se fasse dans le cadre d'une réunion d'un comité qui a l'avantage de permettre l'échange d'opinions contradictoires et la comparaison entre différents chercheurs d'un domaine donné. Cette évaluation doit bien sûr intégrer les indicateurs quantitatifs mais aussi pleinement prendre en compte l'originalité de la recherche et sa pertinence.

L'évaluation par les pairs doit être fondée sur un dossier dans lequel sont présentés les travaux de l'enseignant-chercheur ou du chercheur, en utilisant des formulaires où une place importante est donnée à la sélection des meilleures publications et à la présentation par le chercheur de sa carrière et de ses projets, comme cela se pratique pour les candidats à une subvention individuelle auprès de l'ERC. Cette « *auto-évaluation* » n'implique pas, bien sûr, que le chercheur donne en quoi que ce soit une opinion sur sa carrière. Il peut être intéressant d'ajouter au dossier, au moins dans certains cas, des lettres d'évaluation personnelles émanant de scientifiques reconnus éventuellement proposés par le candidat, comme cela est réalisé avec beaucoup d'efficacité à l'Institut Universitaire de France. Enfin, il faut garder en mémoire que l'évaluation des individus doit s'effectuer en tenant compte du contexte scientifique dans lequel ils sont actifs.

3. Les critères d'évaluation des activités d'enseignement

Jusqu'à présent, l'évaluation des enseignants-chercheurs était presque exclusivement réalisée sur leur activité de recherche, tant pour ce qui concerne le recrutement que les promotions. Mais il convient à l'évidence, et les nouveaux textes le stipulent, d'y associer l'évaluation des deux autres grands types d'activité des enseignants-chercheurs, à savoir l'enseignement et les travaux d'intérêt collectif, comme cela se pratique dans plusieurs pays étrangers comparables au nôtre.

3. A - L'enseignant-chercheur et l'équipe pédagogique

L'évaluation de l'activité des enseignants-chercheurs est compliquée par le fait que cette activité doit se développer en étroite association avec les autres enseignants de la même université, rendant parfois difficile l'évaluation individuelle. Elle est également compliquée par la tendance heureuse à diversifier les formes d'enseignement en donnant une place privilégiée aux interactions directes avec les étudiants, y compris sous la forme d'un tutorat. Ces deux points, qui sont d'ailleurs largement repris dans le décret du 23 avril 2009, sont apparus suffisamment importants pour en rappeler les termes avant de discuter plus avant les modalités d'évaluation des enseignants-chercheurs qui s'inscrivent dans leur contexte.

L'Académie a déjà signalé dans un précédent rapport (*Attractivité des Carrières de Recherche* - 2008) combien la charge statutaire d'enseignement est excessive et n'autorise pas, en particulier un jeune Maître de Conférences, à installer et développer une recherche significative, qu'il appartienne ou non à une équipe labellisée³. Une réflexion doit être

³ Il existe plusieurs catégories d'enseignants à l'université comme p.e. les PRAG et les MdC. Les PRAG, professeurs agrégés, enseignent sur la base d'un double service de MdC (192H eq TP/an x 2) et ne sont pas tenus de faire de la recherche. Les MdC ont une charge statutaire lourde (192H eq TP/an) mais ils doivent faire de la recherche. A l'échelle de l'Etablissement, il faut donc veiller aux cas, rares on peut l'espérer mais

conduite et des décisions prises à ce sujet à l'échelle de chaque établissement. Cela ne veut pas dire pour autant qu'il faille augmenter les tâches d'enseignement des enseignants-chercheurs peu ou pas impliqués dans la recherche. La mise en place d'équipes enseignantes dans chaque établissement, par niveau et par discipline, est indispensable et permet de répondre à cette difficulté. En effet, chaque équipe doit coordonner les enseignements, définir leurs niveaux et les objectifs à atteindre, tout en favorisant les échanges nécessaires à une bonne émulation en son sein.

Un point important concerne la formation des enseignants-chercheurs. Celle-ci ne peut s'envisager qu'au sein de l'équipe pédagogique à laquelle le jeune Maître de Conférences doit obligatoirement être intégré. Une aide à la formation par le biais d'un tutorat d'une à deux années assuré par un enseignant confirmé et reconnu par ses pairs, semble indispensable, notamment pour les jeunes Maîtres de Conférences inexpérimentés. Les expériences antérieures avec les ATER ou les moniteurs n'ont pas toujours été réalisées dans de bonnes conditions de formation à l'enseignement ; ces derniers ont trop souvent été mis en situation d'enseignement sans formation ni travail en équipe, au minimum en binôme avec un maître de stage qualifié. Il convient de rappeler que les CIES ont été créés, il y a une vingtaine d'années pour assurer un minimum de formation aux moniteurs. C'est un point important à souligner, compte tenu du nombre significatif d'allocations de recherche couplées avec un poste de moniteur. Un encadrement par le biais d'un tutorat ne peut qu'être bénéfique tant pour le jeune enseignant que pour les étudiants.

3. B - Les expériences pédagogiques

Des expériences pédagogiques conduites dans certains établissements montrent tout l'intérêt qu'il y a à faire participer les chercheurs à certaines activités d'enseignement. Ces derniers sont souvent demandeurs d'un contact direct avec les étudiants qu'ils peuvent aisément avoir dans ce cadre. Par cette action volontaire, ils déchargent également les enseignants-chercheurs d'une partie de leur lourde charge de service⁴ tout en améliorant leur culture générale de la discipline, culture nécessaire à une bonne recherche. L'idéal, qui a été expérimenté avec bonheur dans quelques établissements, est de confier la responsabilité totale d'un enseignement (niveaux L3 ou M1) à une équipe de recherche (ex : une UMR), en partenariat avec l'équipe pédagogique responsable du niveau concerné, sous la responsabilité de l'UFR ou du département d'enseignement correspondant. D'autres formules sont possibles et méritent d'être expérimentées.

qui existent néanmoins, de MdC (ou de chercheurs) qui ne font pas (ou plus) de recherche et qui ont seulement un service statutaire normal à effectuer dans la même discipline qu'un PRAG. C'est un autre paradoxe du système français qui place côte à côte des personnels de statut différent pour effectuer le même type de travail !

⁴ Les Organismes et Établissements doivent veiller à ce que, dans ces cas précis, la décharge de service effectuée n'entraîne pas la suppression de la prime d'encadrement doctoral et de recherche (PEDR), seul élément discriminant reconnaissant la qualité de l'activité de recherche de certains enseignants-chercheurs.

3. C - L'évaluation de l'enseignement

L'évaluation des activités d'enseignement peut alors être réalisée selon plusieurs approches conduisant à la production d'une grille d'évaluation de l'enseignement aux niveaux national et local, en sachant que le niveau local est évidemment le plus pertinent. Un critère très important d'évaluation reste l'appréciation des étudiants, même si celle-ci peut exposer à des biais : certains enseignants sont bien appréciés des étudiants sans que ceux-ci puissent réaliser les faiblesses du contenu pédagogique ; un professeur peut faire des cours d'une grande clarté, mais présentant des données ou des concepts obsolètes, ce que l'étudiant n'est pas en mesure de discerner. Néanmoins, il est urgent de réfléchir aux modalités de cette évaluation pour l'ensemble des universités en faisant la part du niveau du cursus (L ou M) et des formes d'enseignement (magistral, TD, TP, enseignement à distance, etc., et même tutorat⁵). Il faut souligner le fait que la qualité pédagogique d'un enseignant n'est généralement connue ou appréciée que de ses étudiants et quelquefois de ses proches collègues. La grille d'évaluation de l'enseignement doit donc être élaborée avec les étudiants en tenant compte de ces différentes modalités, et les conclusions doivent être prises en compte par l'enseignant ou l'équipe enseignante (après les examens : réunion de fin d'année, p.ex.) pour améliorer et/ou modifier les enseignements et pour limiter les biais possibles de cette évaluation. L'évaluation des professeurs par les étudiants est très largement pratiquée par les pays anglo-saxons et dans certains établissements français, en particulier les grandes écoles, avec beaucoup d'intérêt et de profit tant pour les enseignés que pour les enseignants⁶. On pourra s'inspirer des pratiques des universités de Delft et de Louvain et des Ecoles polytechniques fédérales de Zurich et de Lausanne. L'équipe enseignante doit répercuter les difficultés locales inhérentes au contexte particulier de son enseignement (hétérogénéité des étudiants, amphithéâtre surchargé et autres conditions matérielles locales).

Les meilleurs Maîtres de Conférences qui s'investissent très fortement en enseignement et qui sont unanimement reconnus pour leurs qualités pédagogiques devraient pouvoir être candidats à des promotions locales, notamment à la hors classe, comme le permet le décret du 23 avril 2009 (art. 22) et comme cela se passe dans certains pays. Il faut récompenser les enseignants exceptionnels par l'octroi d'une prime particulière : sorte de « *teaching award* » à l'échelle de chaque Établissement, élément tangible susceptible de créer une réelle compétition entre les enseignants. Il convient également de valoriser les enseignements les plus élémentaires (ceux de 1^{ère} année, en particulier) et les plus technologiques (IUT, licences professionnelles, p.ex.). A l'opposé, des enseignants-chercheurs (MdC ou PR) qui délibérément négligent leurs activités d'enseignement (absence fréquente, qualité de l'enseignement discutable, etc.) devraient en subir les conséquences. Dans cet esprit, l'avis des étudiants doit pouvoir être

⁵ Une incitation au suivi et à l'aide personnalisée des étudiants doit être mise en place par les équipes enseignantes sous une forme à définir (tutorat, rencontres avec les enseignants sur des créneaux horaires fixés en début d'année universitaire : exemple des 'Office Hours', etc.)

⁶ Une aide pédagogique par une équipe spécialisée devrait pouvoir être apportée à l'enseignant en cas de difficulté.

relayé par l'équipe enseignante locale et pris en compte lors de l'examen des promotions à l'échelle nationale. Mais pour cela, il faut comme il est dit plus haut, rendre les critères d'évaluation de l'enseignement effectué dans le cadre de la formation initiale ou continue, plus objectifs que ceux actuellement utilisés, et adaptés si besoin est, discipline par discipline.

L'évaluation des activités d'enseignement doit évidemment intégrer certains critères objectifs comme le contenu et l'originalité de l'enseignement (publications à caractère pédagogique, ouvrages édités, mise en ligne de cours ou TP-TD, diffusion à l'étranger et traduction de ces ouvrages, expositions, etc.). Cette évaluation pourrait s'appuyer sur les archives pédagogiques existantes qu'il conviendrait d'actualiser en recensant notamment de façon plus systématique les ouvrages ou documents pédagogiques produits par les enseignants.

Pour l'information des évaluateurs (experts nationaux ou locaux), un relevé des charges annuelles d'enseignement doit, selon les textes officiels, être publié en toute transparence chaque année par l'établissement, les éventuelles décharges d'enseignement devant être clairement mentionnées et justifiées.

Enfin, le pourcentage de succès, le devenir des étudiants par niveau d'étude et par discipline (ex : L1, L2, L3, M1, M2, doctorat) et les débouchés des étudiants par filière et à l'issue de la thèse, doivent être rendus publics et pris en compte pour l'évaluation individuelle des enseignant-chercheurs, même s'il ne sera pas toujours facile d'identifier ce qui revient à chaque enseignant.

4. L'évaluation des activités d'intérêt collectif

La Loi relative aux libertés et responsabilités des universités innove en distinguant 6 missions de service public de l'enseignement supérieur. Outre les missions traditionnelles concernant la formation initiale et continue, et la recherche scientifique et technologique avec la diffusion et la valorisation de ses résultats, la Loi ajoute 4 missions complémentaires que la collectivité universitaire doit assurer, à savoir : l'orientation et l'insertion professionnelle, la diffusion de la culture et l'information scientifique et technique, la participation à la construction de l'espace européen de l'enseignement supérieur et de la recherche, et la coopération internationale. Ces 4 dernières missions relèvent toutes d'actions d'intérêt collectif. Le nouveau décret du 23 avril 2009 fixant les dispositions statutaires applicables aux enseignants-chercheurs confirme que l'ensemble de leurs activités doit être pris en considération, et il en résulte que les résultats obtenus dans chacune d'elles par les enseignants-chercheurs devront donc être évalués, notamment les charges administratives et/ou collectives qui impliquent de véritables responsabilités (coordonnateur de 1^{ère} année L1, responsabilité d'UFR, direction de la recherche au sein de l'établissement, direction d'une Ecole Doctorale, d'un Institut Fédératif de Recherche, relations internationales, etc.). L'organisation de ces charges ne doit cependant pas être trop consommatrice de temps et

d'énergie de manière à ce qu'il n'y ait pas, comme c'est le cas aujourd'hui, une véritable inflation de réunions qui stérilisent nombre de chercheurs et d'enseignants-chercheurs ou qui leur fournissent un alibi pour ne pas accomplir leurs missions fondamentales, en particulier la recherche.

Parmi les travaux d'intérêt collectif qui sont désormais pris en considération, il convient de mentionner les charges de direction d'établissements, d'organismes, d'unités de recherche, de conseil, de participation aux jurys d'examen et de concours, de tutorat, d'orientation des étudiants, d'aide à leur insertion professionnelle, de coopération avec les entreprises publiques ou privées, de développement, expertise et coordination de la recherche fondamentale, appliquée, pédagogique et technologique, ainsi que de la valorisation de ses résultats. De plus, la nécessité de mieux expliciter vis-à-vis de la société la nature et les résultats de la science, produit une diversité d'initiatives et de réalisations qui doivent être reconnues, comme le prévoit le décret du 23 avril 2009 selon lequel les enseignants-chercheurs « *contribuent au dialogue entre sciences et sociétés notamment par la diffusion de la culture et de l'information scientifique et technique*. Enfin, la fonction d'évaluateur devrait être prise en compte dans le cadre des actions d'intérêt collectif.

Un relevé de ces activités devrait être régulièrement publié par l'établissement pour chaque enseignant-chercheur.

L'évaluation de ces actions d'intérêt collectif est devenue une nécessité dans les nouveaux textes en vigueur, mais les outils disponibles ne sont pas explicités. Il y a donc là un chantier nouveau à explorer qui pourrait s'inspirer des manières d'évaluer d'autres activités collectives et administratives de l'État.

5. La fréquence et la forme des évaluations

Il convient de distinguer les évaluations approfondies destinées aux recrutements, promotions, changements de grade ou de classe d'une part, et les évaluations de suivi d'activité d'autre part.

Pour les premières, il faut en réduire le nombre et la fréquence aux seules étapes importantes de la carrière d'un chercheur ou d'un enseignant-chercheur : les recrutements⁷, les concours ou les promotions de grade ou changements de corps : passage CR-DR2 ou MdC-PR2, DR2-DR1 ou PR2-PR1, l'accès à la classe exceptionnelle ou encore à la HC des MdC.

⁷ Les Établissements doivent afficher leurs orientations scientifiques et pédagogiques de manière à ce que les appels à candidatures ne soient pas biaisés. Les candidats retenus pour audition doivent avoir leurs frais de déplacement couverts par l'Établissement. Les taux de recrutement extérieur doivent figurer dans le rapport quadriennal ainsi que l'origine des MdC. L'auto-recrutement au sein de l'Établissement a été dramatique dans certaines universités.

Il convient aussi de limiter le nombre des évaluations de suivi en les faisant coïncider avec le contrat d'établissement (c'est-à-dire tous les 4 ans), évaluations effectuées sur la base de fiches d'activité, ces dernières étant simplement destinées au suivi normal de l'activité des personnels. Actuellement, ces fiches existent et sont utilisées par les enseignants-chercheurs qui effectuent leur recherche au sein d'unités labellisées (ex : UMR CNRS). Une procédure particulière est utilisée par les EPST ou les EPIC pour leurs chercheurs statutaires. Par exemple au CNRS, il existe des fiches annuelles d'activité qui sont à remplir de manière statutaire par les chercheurs. Leur examen par les sections compétentes du Comité national de la recherche scientifique, indépendamment de l'évaluation des laboratoires et des équipes de recherche, est une procédure longue, inutile, coûteuse en temps et en personnels scientifiques et administratifs. Une évaluation de suivi d'activité à 4 ans, lors du quadriennal recherche, est amplement suffisante. Ces fiches ont pour objectif de permettre au président d'université ou au directeur d'organisme d'identifier les enseignants-chercheurs ou les chercheurs en difficulté. Il leur reviendrait donc, dans ce cas, de demander un rapport approfondi à des experts qu'ils désigneront, ce qui permettra de préparer un entretien éventuel dans le cadre du dispositif défini plus bas (voir § 7).

En tout état de cause, toute évaluation devrait se terminer par un document écrit qui permettrait de garder l'historique et la mémoire.

Par ailleurs, l'Académie souhaite faire des commentaires et des recommandations sur deux étapes actuelles de la carrière des enseignants-chercheurs dont l'évaluation aboutit à une inscription sur une des deux listes de qualification, soit aux fonctions de Maître de Conférences, soit aux fonctions de Professeur. Ces listes d'aptitude ont été créées pour pallier l'hétérogénéité de niveau des thèses et des HDR. L'autonomie des établissements devrait conduire très rapidement à une importante redéfinition, au niveau local, des critères nécessaires à l'obtention de la thèse et de l'HDR et aux rôles essentiels que doivent jouer les écoles doctorales d'une part et les conseils scientifiques de l'autre. De plus, la responsabilité des évaluateurs dans les jurys de thèse et d'HDR devrait être largement prise en compte dans l'évaluation de leur dossier personnel, le nombre d'appartenances à des jurys, critères figurant souvent dans les dossiers d'évaluation individuelle, devant être pondéré par le niveau de l'implication des membres du jury, au demeurant difficile à évaluer (cf. § 6). Ces recommandations devraient améliorer le niveau de reconnaissance des HDR et des thèses. Cependant, la notion de diplôme national, chère au système égalitaire français, est antinomique de celle d'autonomie des universités. Les conséquences sont doubles : a) un certain discrédit de la thèse et de l'HDR par les employeurs du secteur privé dont les dirigeants méconnaissent souvent le contenu et la qualité ; b) une charge de travail stérile et considérable pour les candidats et les collègues chargés de l'évaluation (CNU). Enfin, il faut souligner l'actuelle lourdeur du système français où quatre étapes d'évaluation liées aux inscriptions préalables sur les listes de qualification sont nécessaires au passage comme

professeur, à comparer aux deux étapes requises dans les pays comparables au nôtre. Une réflexion globale à ce sujet doit être entreprise.

6. Les évaluateurs

6. A - Le choix des évaluateurs

La qualité des évaluateurs est fondamentale. Il faut à tout prix éviter qu'à tous les niveaux de l'évaluation, l'évaluateur n'ait pas la compétence nécessaire, voire une compétence inférieure à celle de la personne évaluée. L'Académie propose, pour chaque catégorie d'évaluateurs, la publication de pré-requis qui pourraient être préparés par l'AERES⁸. Chaque évaluateur devrait fournir un court CV avec, par exemple, 10 publications (ou travaux d'application) publiés au cours des 10 dernières années (nombre à ajuster éventuellement pour les jeunes recrutés).

Afin de garantir compétence et excellence scientifiques, tous les évaluateurs devront être évalués, que ce soit au niveau de l'Établissement ou au niveau national (CNU, CNECA, ...).

Une des difficultés souvent rencontrée est la résistance des meilleurs enseignants-chercheurs ou chercheurs à accepter de participer aux comités d'évaluation par manque de temps ou d'intérêt. Le problème du manque de temps, d'autant plus crucial que l'on souhaite faire participer les meilleurs, peut être partiellement contourné en réduisant à trois ans la durée du mandat des évaluateurs, mesure déjà proposée pour des raisons de déontologie mentionnées plus haut. S'agissant de la motivation, la question est plus difficile, mais la mise en valeur de cette fonction d'évaluation pour la carrière pourrait être un élément d'attractivité.

La question est par ailleurs souvent posée de savoir s'il est opportun d'indemniser les évaluateurs et à quel niveau. Il ne paraît pas nécessaire de le faire pour les évaluateurs de nationalité française, dans la mesure où l'ensemble des frais de séjour et de voyage engagés par les évaluateurs qui doivent se déplacer serait correctement pris en charge. Certaines indemnités existent déjà, notamment pour les concours au CNRS et d'autres seront mises en place pour le CNU par application du décret du 23 avril 2009 modifiant les dispositions relatives au Conseil des Universités. Le problème est plus délicat pour les évaluateurs étrangers à qui on demande un travail parfois important au service d'organismes de recherche qui ne les concernent pas. Les indemnités pour les évaluateurs étrangers sont sans doute utiles mais doivent rester d'un montant raisonnable. Elles doivent surtout être accompagnées de meilleures conditions d'accueil lorsque les étrangers sont invités à siéger en France, notamment par les remboursements de frais de séjour et de voyage.

⁸ En Allemagne, par exemple, les évaluateurs pour un projet de type ANR sont élus par les jeunes chercheurs ayant passé leur thèse depuis plus de trois ans.

6. B - La place respective de l'évaluation nationale et de l'évaluation locale

S'agissant du rôle respectif de l'évaluation nationale et de l'évaluation locale, il apparaît clairement que l'activité scientifique doit être évaluée au niveau national (CNU, organismes de recherche, ou autre commission nationale) - à défaut de pouvoir l'être à un niveau européen qui garantirait un pool plus large et plus indépendant d'évaluateurs - même si les instances locales doivent pouvoir exprimer leur opinion. En revanche, pour ce qui concerne les activités d'enseignement et les travaux d'intérêt collectif, l'essentiel des actes d'évaluation doit être réalisé au niveau local.

Cette dualité d'évaluation pose le problème des comités d'évaluation propres à chaque université qui doivent être gérés localement mais inclure une fraction importante de membres extérieurs à l'université (la LRU - art. 25, stipule 50% pour les comités de recrutement), le terme d'extérieur incluant à la fois des experts français n'appartenant pas à la même université ou au même groupement (par exemple le même pôle universitaire), et des membres étrangers.

Les consignes aux évaluateurs devront être clairement explicitées dans chaque Établissement et organisme.

7. L'utilisation de l'évaluation

Une des difficultés majeures du système d'évaluation est l'absence fréquente de conséquences de l'évaluation, une fois le chercheur ou l'enseignant-chercheur recruté. Il faut donc veiller à ce que l'évaluation soit de préférence réalisée lorsqu'elle peut conduire à une promotion ou à une aide à une réorientation, ou encore à une autre forme tangible de reconnaissance (par exemple une prime).

7. A - Les recrutements

Un point capital de l'évaluation, où l'erreur devrait être l'exception, est le recrutement, compte tenu du statut de fonctionnaire d'État des chercheurs et des enseignants-chercheurs. Pour cela, l'organisme de recherche ou l'Établissement universitaire doivent se doter de procédures transparentes et d'un panel d'évaluateurs compétents (cf. § 1 et 6). Par ailleurs, les affichages de postes et les modalités pratiques de candidature doivent être publiés suffisamment à l'avance, notamment pour les candidats étrangers ou les candidats français en séjour à l'étranger. Enfin, il faut généraliser un système à deux tours avec une admissibilité sur dossiers et une admission à l'issue d'une présentation orale significative et d'un entretien approfondi. Les critères de créativité et d'originalité des recherches requis pour un

recrutement étant difficiles à apprécier par le simple jeu d'une analyse bibliométrique ou d'une audition préparée à l'avance, l'examen approfondi par les pairs devrait jouer pleinement ce rôle. En effet, la caricature de présentation et d'entretien (10 à 30 min pour tous les candidats) qui a lieu dans certains des grands organismes de recherche, en raison du trop grand nombre de candidats auditionnés, est largement responsable d'erreurs de recrutement.

Par ailleurs, les Établissements devraient participer au remboursement des frais liés au déplacement des candidats admissibles, pour ne pas limiter la candidature de certains jeunes aux faibles ressources, et responsabiliser un peu plus les Commissions de recrutement des Établissements. Enfin, il est important de rappeler qu'un pourcentage très significatif d'évaluateurs extérieurs à l'Établissement universitaire doit participer au Comité de recrutement (la loi LRU stipule 50%) pour éviter les recrutements internes très fréquents dans certaines universités.

7. B - Les promotions

Pour les promotions importantes ou lors de changements de corps (passages CR-DR, ou MdC-PR) ou encore dans le cas de demande de mutation, les mêmes règles de transparence, de compétence et d'éthique doivent s'appliquer (cf. §1 et § 5) ainsi que les pourcentages d'évaluateurs extérieurs.

7. C - Le suivi de carrières

Enfin, se pose le problème très important de l'évaluation qui participe au suivi régulier du déroulement de carrière. Dans la plupart des cas, l'organisme de recherche ou l'Établissement universitaire transmet le rapport du comité d'évaluation aux chercheurs ou enseignants-chercheurs évalués pour les informer et saluer l'excellence, la pertinence ou la qualité du travail réalisé. Des préconisations particulières peuvent aussi être formulées par les responsables des organismes ou des Établissements universitaires pour optimiser le travail des personnels évalués. Cependant, comme le système français actuel de fonctionariat ne permet pas de revenir sur un recrutement qui se révèle inadapté à l'intéressé, il faut aménager la possibilité d'une évolution tout au long de la carrière pour les personnels recrutés. Il convient donc de prendre en compte les modifications de leur trajectoire professionnelle, ou l'évolution de leurs profils de compétence, ainsi que leurs souhaits personnels pour ne pas laisser s'installer des situations de démissions ou de 'mal-être' relativement importants qui n'ont, aujourd'hui, aucune solution satisfaisante. Une possibilité serait que les enseignants-chercheurs puissent se prononcer eux-mêmes à intervalles réguliers (4 ans) sur l'orientation qu'ils souhaiteraient privilégier, notamment sur le partage de leur temps entre recherche et enseignement. Ils seraient ensuite évalués sur l'orientation choisie.

7. D - La mise en place d'un dispositif de suivi des carrières

Quand les conclusions de l'évaluation posent problème, un Comité ad-hoc, placé sous l'autorité d'une personnalité scientifique reconnue pour ses qualités humaines et de management, devrait avoir un entretien approfondi avec l'intéressé pour l'aider à trouver la meilleure solution à ses difficultés (localement, au sein de l'Établissement ou dans d'autres ministères de la fonction publique, avec une formation appropriée si nécessaire). Chaque organisme et chaque université doivent se doter de tels Comités et faire le recensement des personnalités capables de jouer ce rôle clé sur la base du volontariat. Un tel dispositif, qui pourrait s'apparenter à une commission de type paritaire, devrait être mis en place rapidement à l'échelle de chaque établissement.

Actuellement, aucune solution n'est recherchée de façon professionnelle. Le système français est paralysé par une regrettable et souvent désastreuse confusion des genres au niveau de l'évaluation avec des Comités qui sont, en réalité, constitués comme des commissions paritaires et où l'on confond souvent compétence, recherche de l'excellence et gestion des cas individuels. La solution à rechercher dans le cas de l'évaluation d'un suivi de carrière problématique est, de facto, laissée à l'initiative individuelle du responsable d'enseignement ou de recherche, à l'échelle de l'Établissement (UFR ou directeur d'une UMR) ; et généralement peu de solutions satisfaisantes sont proposées à l'intéressé évalué. Cela conduit à un véritable gâchis de compétences et de moyens humains alors que l'essentiel des forces du dispositif d'enseignement et de recherche s'appuie sur ses ressources humaines. Il faut donc pallier cette carence du système français par un dispositif à expérimenter sans pour autant nécessairement créer une structure administrative supplémentaire.

CONCLUSION

L'évaluation est un processus incontournable pour assurer un fonctionnement moderne et efficace de l'activité des chercheurs et des enseignants-chercheurs dans toutes les missions qui leur sont dévolues, recherche, enseignement et activités d'intérêt collectif, en sachant que la part d'importance de chacune de ces missions est susceptible d'évoluer au cours de la carrière. Ce processus complexe et sensible de l'évaluation demande tout à la fois un grand professionnalisme, un entourage éthique rigoureux, et une acceptation consensuelle des « règles du jeu ».

Outil majeur de direction et d'animation des organismes de recherche et des universités, l'évaluation est désormais reconnue et acceptée. Mais au-delà de ce contrôle indispensable pour les instances décisionnelles, il faut la considérer comme partie intégrante du processus de production de la recherche, de l'enseignement, et des actions collectives associées. En

effet, elle doit aussi avoir pour finalité profonde d'aider les chercheurs et enseignants-chercheurs à progresser en vue d'atteindre ou de maintenir un niveau d'excellence dans leur domaine. Ce regard extérieur compétent peut être l'occasion, assez rare, de faire le point sur l'orientation de leurs travaux et éventuellement si nécessaire, de fournir une aide à une réorientation, qu'il s'agisse de thèmes de recherche ou de carrière. Pour ces raisons, elle doit être considérée comme une chance par les chercheurs et les enseignants-chercheurs.

Les problèmes de l'évaluation se sont posés dans tous les grands pays scientifiques. Les solutions diffèrent selon les cultures et les statuts des personnels. Il conviendrait que l'organisation de l'évaluation en France s'inscrive au niveau européen, aussi bien par l'adoption de règles communes que par les échanges d'évaluateurs entre les différents pays.

ANNEXE 1
COMPOSITION DU GROUPE DE TRAVAIL

Jean-François BACH (coordonnateur général)
Secrétaire Perpétuel de l'Académie des Sciences
Professeur à l'Université René Descartes

Christian DUMAS (co-président)
Professeur à l'Ecole Normale Supérieure de Lyon

Denis JÉROME (co-président)
Directeur de recherche émérite au CNRS

Yves AGID
Professeur à l'Université Pierre et Marie Curie
Praticien hospitalier à l'Hôpital de la Pitié Salpêtrière

François BACCELLI
Directeur de recherche à l'INRIA

Gérard BERRY
Professeur au Collège de France
Président de la Commission Évaluation de l'INRIA

Pierre BRAUNSTEIN
Directeur de Recherche au CNRS

Alain CARPENTIER
Vice-Président de l'Académie des Sciences
Professeur émérite à l'Université Paris Descartes

Catherine CESARSKY
Haut Commissaire à l'Énergie Atomique

Bruno CHAUDRET
Directeur de Recherche au CNRS

Vincent COURTILOT
Directeur de l'Institut de Physique du Globe à Paris
Professeur à l'Université Denis Diderot

Jean DERCOURT
Secrétaire Perpétuel de l'Académie des Sciences
Professeur émérite à l'Université Pierre et Marie Curie

Bernard DUJON
Professeur à l'Université Pierre et Marie Curie
Professeur à l'Institut Pasteur

Marc FONTECAVE
Professeur au Collège de France

Jacques FRIEDEL
Professeur émérite à l'Université Paris-sud Orsay

Michel LE MOAL
Professeur émérite à l'Université Victor Ségalen de Bordeaux

Ghislain de MARSILY
Professeur émérite à l'Université Pierre et Marie Curie

Bernard MEUNIER
Président-directeur général de PALUMED

René MOREAU
Professeur émérite à l'Institut national polytechnique de Grenoble

Eva PEBAY-PEYROULA
Directrice de l'Institut de biologie structurale

Olivier PIRONNEAU
Professeur à l'Université Pierre et Marie Curie

Jean-Pierre RAMIS
Professeur à l'Université Paul Sabatier de Toulouse

Jean SALENÇON
Président de l'Académie des Sciences
Professeur honoraire à l'École Polytechnique et à l'École nationale des Ponts et Chaussées

André SENTENAC
Conseiller scientifique à la Direction des sciences du vivant
Commissariat à l'Énergie Atomique

Alain-Jacques VALLERON
Professeur à l'Université Pierre et Marie Curie

Jean-Christophe YOCCOZ
Professeur au Collège de France

Coordination de la rédaction :
Brigitte d'ARTEMARE
Chef de Cabinet des Secrétaires Perpétuels de l'Académie

ANNEXE 2
LISTE DES PERSONNALITÉS AUDITIONNÉES
DU 16 MARS AU 30 JUIN 2009

Serge BAUIN

Directeur de l'Unité Propre de Service *Indicateurs, Programmation, Allocations des Moyens*
du CNRS

Lionel COLLET,

Président de la Conférence des Présidents d'Université (CPU)

Jean-François DHAINAUT, Président

Pierre GLORIEUX, Directeur de la section Recherche

Agence d'Evaluation de la Recherche et de l'Enseignement Supérieur (AERES)

Ghislaine FILLIATREAU

Directrice de l'Observatoire des Sciences et des Techniques (OST)

Franck LALÖE

Directeur de Recherche Émérite au Laboratoire Kastler-Brossel de l'École Normale
Supérieure de Paris (ENS)

Jean-Claude LEHMANN

Professeur Honoraire de l'Université Pierre et Marie Curie

Ancien Président de l'Académie des Technologies

ANNEXE 3

GLOSSAIRE

AERES	Agence d'Évaluation de la Recherche et de l'Enseignement Supérieur
ANR	Agence Nationale de la Recherche
ATER	Attaché Temporaire d'Enseignement et de Recherche
CIES	Centre d'Initiation à l'Enseignement Supérieur
CNECA	Commission Nationale des Enseignants-Chercheurs dépendant du ministère de l'Agriculture
CNRS	Centre National de la Recherche Scientifique
CNU	Conseil National des Universités
EPIC	Établissement Public à caractère Industriel et Commercial
EPST	Établissement Public à caractère Scientifique et Technologique
ERC	European Research Council
HDR	Habilitation à Diriger des Recherches
IF	Impact Factor
INIST	Institut de l'Information Scientifique et Technique
INRIA	Institut National de Recherche en Informatique et en Automatique
IUF	Institut Universitaire de France
IUT	Institut Universitaire de Technologie
LRU	Loi relative aux Libertés et Responsabilités des Universités
OST	Observatoire des Sciences et des Techniques
PRAG	Professeur Agrégé
UFR	Unité de Formation et de Recherche
UMR	Unité Mixte de Recherche

ANNEXE A : OÙ TROUVER L'INFORMATION DE BASE ?

Deux bases de données majeures existent actuellement, ISI-Thomson (Reuters) et SCOPUS (Elsevier). Toutes les deux sont payantes mais les chercheurs du CNRS disposent d'un accès à ISI via l'INIST. Il serait bon qu'un accès analogue soit accordé aux membres de l'enseignement supérieur. Des études ont montré que la base ISI bien que plus modeste que SCOPUS rassemble l'essentiel des données bibliographiques depuis le début du vingtième siècle⁹. La base ISI ne cesse de s'améliorer, notamment par la prise en compte des actes de colloques et des ouvrages. Elle représente donc une base sérieuse à partir de laquelle l'information peut être traitée avec des outils bibliométriques appropriés. Dans le cas de la discipline des Mathématiques la base Math Reviews est beaucoup plus informative que les bases généralistes citées plus haut.

⁹ SCOPUS couvre un nombre de revues considérablement plus élevé que ISI (16000 contre 9000). Cependant, ISI a rétro-indexé ses revues jusqu'au début du 20^{ème} siècle. Des études ont montré que la base ISI couvre 80 à 100% des citations pour les domaines de biochimie, biologie moléculaire, chimie et physique. La couverture n'est que de 40 à 60% pour les mathématiques et les sciences pour l'ingénieur. Il est à noter que les actes de congrès et les ouvrages sont de mieux en mieux pris en compte par la base ISI.

ANNEXE B

DES INDICATEURS À ÉVITER

Le facteur d'impact d'une revue pour l'année N est défini par la base de données ISI par le rapport entre le total du nombre de citations vers cette revue à partir de toutes les revues enregistrées dans la base ISI durant les années N-1 et N-2 arbitraire et la somme des articles publiés par la revue durant ces mêmes années.

Le facteur d'impact n'est pas adapté à l'évaluation car il peut masquer la réalité. En effet, des magazines pluridisciplinaires comme Nature ou Science possèdent un facteur d'impact de l'ordre de 28-30 alors qu'il est plutôt de l'ordre de 3-6, en moyenne¹⁰, pour les journaux spécialisés de sociétés savantes. Le facteur d'impact d'une revue équivaut sommairement à un indice de citation annuel moyen des articles de la revue. Cependant, il faut garder en mémoire qu'une très petite fraction d'articles bénéficie d'un taux de citations élevé¹¹. En conséquence, le taux moyen de citations peut fort bien, suivant les domaines, s'éloigner fortement du taux moyen d'un domaine en particulier.

En résumé, l'utilisation du facteur d'impact dans un but d'évaluation d'un niveau scientifique individuel est sans fondement et doit être proscrite. De plus, il faut souligner que la prise en compte exagérée dans l'évaluation d'un chercheur du facteur d'impact de la revue dans laquelle il publie est dangereuse, car cela condamne à terme toutes les revues scientifiques qui ne sont pas au plus haut niveau, à disparaître.

¹⁰ Ces valeurs sont très largement dépendantes de chaque discipline ou sous-discipline.

¹¹ Dans un journal classique environ 15% des articles génère la moitié des citations à la revue. A titre indicatif, pour une revue de prestige telle que Nature, 77% des articles n'atteignent pas pendant la période de deux ans, le taux moyen de citations qui est déterminé par le facteur d'impact. Cette proportion atteint même 80% pour la revue Science.

ANNEXE C

INDICATEURS DE PRODUCTIVITÉ

J. Hirsch de l'Université de Californie a suggéré en 2005 une mesure de la qualité scientifique par l'indice h ¹². Cet indicateur est devenu rapidement très populaire et tend même à devenir un outil classique que les chercheurs utilisent sans trop connaître ses limites et défauts. Il en est de même pour des administrations qui viennent de s'en emparer.

L'indicateur h est déterminé en classant les articles d'un auteur en fonction du nombre décroissant, C_i , de leurs citations, de sorte que le papier de rang $i=h$ ait reçu au minimum h citations. Cet indicateur est un nombre associé à un chercheur et est calculé en considérant l'impact de l'ensemble de ses publications (mais il peut également être étendu à d'autres ensembles comme un groupe de chercheurs, un laboratoire, un domaine, un pays, des journaux, etc). Il tient compte de tous les papiers très cités sans toutefois donner un avantage particulier à ceux qui ont été abondamment cités et dont l'évolution de la science ne peut se passer. Cependant, il présente l'intérêt d'éliminer de l'évaluation les très (et souvent trop) nombreux articles d'un auteur qui sont peu ou même jamais cités. De nombreuses variantes normalisées de l'indicateur h ont été développées tenant compte de la séniorité des chercheurs ou du nombre d'années d'activité professionnelle comme p. ex. le facteur $m=h/\text{nombre d'années depuis la première publication}$.

Ce qui a rapidement contribué à la grande popularité de cet indicateur (parmi les chercheurs) est sa facilité de calcul à partir des bases de données bibliographiques. Il est même calculé en un clic avec le logiciel *Web of Science* de la base ISI-Thomson. Toutefois, cette apparente facilité peut être trompeuse pour un utilisateur non averti. Comme pour tous les indicateurs bibliométriques le problème majeur laissé à la charge de l'évaluateur réside : dans la pertinence du fichier bibliographique de départ. Ce travail préliminaire souvent indispensable doit être effectué par des experts : le bon choix des domaines scientifiques, la suppression des homonymies, le problème des noms composés¹³, etc.

¹² J.E. Hirsch *An index to quantify an individual research output*, Proceedings of the National Academy of Sciences, 102, 16569-16572, 2005

¹³ En outre, il existe plusieurs manières d'utiliser la base de données ISI Web of Knowledge. La première (et la plus aisée) consiste à employer l'option Advanced Search de Web of Science qui permet d'obtenir pour un auteur donné la liste des articles dans les journaux répertoriés par la base. La seconde, plus délicate, utilise l'option Cited Reference Search qui fournit pour un auteur donné la liste de tous ses articles publiés, même lorsqu'ils sont apparus dans des revues non répertoriées par la base ou des articles publiés dans des actes de congrès. Comme une absence dans la base est aussi très souvent due à des négligences des auteurs citant ou à des erreurs typographiques de références, la première option peut conduire à une minoration significative des taux de citations.

Cependant, l'indicateur h souffre de nombreuses insuffisances. Il ne permet pas de déceler l'excellence dans la grande productivité. Il n'est pas adapté, par exemple, au chercheur qui n'aurait publié qu'un nombre restreint d'articles tout en ayant attiré l'attention de nombreux autres auteurs puisqu'il ne donne pas d'avantages aux articles cités un très grand nombre de fois. Ainsi, l'indicateur h ne permet pas de discriminer entre deux profils de chercheurs de même indicateur : l'un, avec un total modeste des citations jusqu'à l'article de rang h , et l'autre présentant un total élevé des citations pour les h premiers articles. En outre, l'indicateur n'évolue pas lorsque ce sont les premiers h articles à nombre de citations $\geq h$ qui voient leurs citations augmenter au cours du temps. C'est souvent le cas, puisque les articles les plus cités jouissent *de facto* d'une plus grande visibilité, une situation qui accroît encore plus la possibilité de les citer, en quelque sorte la manifestation d'un phénomène d'avantage cumulé.

ANNEXE D

INDICATEUR D'EXCELLENCE DANS LA PRODUCTION

Certains des défauts de l'indicateur de Hirsch ont toutefois été corrigés par L.Egghe qui a proposé, en 2006, un indicateur g visant à privilégier la qualité des travaux scientifiques plutôt que leur quantité¹⁴ et ¹⁵. Les articles étant classés par ordre de nombre de citations C_i décroissant, l'article (réel ou virtuel) de rang g est tel que le total des citations reçues par les g premiers articles soit supérieur ou égal au carré de g ¹⁶. L'indicateur g qui est déterminé par une intégrale des citations rend beaucoup mieux justice que ne le fait h à l'impact de travaux fortement cités même s'ils ne sont qu'en nombre modeste puisqu'il n'est pas comme l'est le facteur h borné par le nombre total de papiers publiés. Cet indicateur met donc plus l'accent sur la qualité. L'indicateur g n'est pas fourni jusqu'à présent par les bases de données ; il demande donc un travail à partir de chaque fichier bibliométrique¹⁷. Comme pour l'indicateur h , il nécessite un travail préliminaire sur la base pour obtenir des fichiers bibliographiques propres.

¹⁴ L. Egghe, *Theory and practice of the g-index*, *Scientometrics*, 69,131, 2006

¹⁵ D. Jérôme, *Redouter l'évaluation*, *Reflets de la Physique*, 10,18,2008

¹⁶ D'une manière plus rigoureuse, les conditions, $\sum_{i=1}^{i=g} C_i \geq g^2$ et $\sum_{i=1}^{i=g+1} C_i < (g+1)^2$ sont satisfaites.

¹⁷ Il est prévu de mettre une aide au calcul des indicateurs sur le site de l'Académie des sciences.

ANNEXE E

BIAIS ET PERVERSIONS

Très tôt, les chercheurs ont tenté d'améliorer leur classement bibliométrique par une manipulation des indicateurs. Par exemple, en publiant en priorité dans des revues de prestige à fort facteur d'impact lorsque ce facteur est pris en compte pour le jugement. Ce comportement peut entraîner une fâcheuse incidence sur la rédaction (formatage de l'article) pouvant même aller jusqu'à mettre en doute l'honnêteté de l'information transmise. Certains indicateurs peuvent inciter des chercheurs à suivre des voies normées plutôt qu'à favoriser leur créativité. Lorsque les citations sont prises en compte pour l'évaluation, le cas de cercles d'auteurs se citant entre eux est assez courant. De même, l'excès d'autocitations ou le tronçonnement d'un long article en plusieurs articles courts (salami slicing) sont aussi des pratiques courantes. Il est assez facile d'identifier ces « professionnels de la publication » par le fait que leurs articles ont souvent un index de citations inférieur au facteur d'impact du journal dans lequel ils sont publiés.

Enfin, certains éditeurs de revues n'hésitent pas à recommander à leurs auteurs de faire référence plus que nécessaire à des papiers publiés dans la même revue dans le seul but d'améliorer le facteur d'impact de la revue.

Selon les disciplines, le nombre d'auteurs varie beaucoup et surtout le choix de l'ordre dans lequel ils figurent est très différent (ordre alphabétique, séniorité ou autre). Ce point peut poser de sérieux problèmes lorsqu'un auteur est cité dans des articles où il n'a pas joué de rôle central. Ce problème devra être analysé avec soins par les pairs.

Autre question importante : doit-on normaliser la citation par le nombre de co-auteurs ? La réponse est loin d'être évidente. Des indicateurs normalisés seront certainement très différents de ceux d'origine surtout pour les domaines dans lesquels les publications admettent un grand nombre de co-auteurs ou lorsque des théoriciens doivent être comparés à des expérimentateurs. Mais les indicateurs obtenus ainsi seraient-ils plus fiables ? Le progrès viendra peut-être d'une pondération de chaque auteur en fonction de sa contribution au travail publié. Pour l'instant, seul une fois de plus l'examen par les pairs permet d'interpréter la qualité de l'information obtenue par la valeur des indicateurs.

Pour les quelques raisons présentées ci-dessus, les indicateurs à développer devront être pensés de manière à déjouer, sinon à minimiser, les biais et les possibilités de perversion.