INSTITUT DE FRANCE
Académie des sciences

INDIVIDUAL EVALUATION OF FULL TIME AND ACADEMIC RESEARCHERS IN THEORETICAL AND EXPERIMENTAL SCIENCES

Report presented on 8 July 2009
to the Minister of Higher Education and Research

CONTENTS

Individual evaluation of full time and academic researchers in theoretical and experimental sciences

Main recommendations

Guidelines for an unquestionable code of ethics

1 – Evaluators must be of high quality and follow the standards set by the national agency for higher education and research evaluation (AERES).
2 – A compulsory declaration of direct and indirect conflicts of interests must be imposed and signed by the members of evaluation committees and, if necessary, replacement within the evaluation committee should be made.
3 – The mandate must be limited to three years, with renewal of a third of the members each year so as to break up lobbies.

Evaluation of research performance

4 – The primordial value of evaluating scientific work directly on the basis of original publications and on eventual interviews must be reaffirmed.
5 – A framework must be set up to ensure a rigorous use of bibliometry. Bibliometry must be restrained to rigorous use. Biases that are often neglected must be avoided, criteria must be adapted to the relevant field and bibliometry should be integrated within a qualitative evaluation. The search for new indicators should be encouraged.
6 – Researchers in applied fields should be given the place they deserve.
7 – Objective criteria based on national and international recognition must be taken into consideration.

Evaluation of teaching performance

8 – An evaluation scale based to a large extent on student appreciation must be established.
9 – Objective criteria such as content and novelty both in teaching and production of educational books and tools must be integrated.

Evaluation of activities benefiting collective interest

10 – Activities that benefit the collective interest must be taken into account, in particular those involving important responsibilities.

Other important measures

11 – The number of evaluations must be kept to a minimum, and a distinction must be made between in-depth evaluations (recruitment and other key career stages) and performance evaluations.

12 – Recruitment should be a two-step process, with a first cut-off based on the dossier, followed by an interview and in-depth examination of successful candidates.

13 – An in-depth evaluation should only be carried out if it will lead to significant consequences for the person examined. Human resources structures must be created to counsel research and teaching staff encountering difficulties.

14 – The role of national evaluations (essentially for research activities) and local evaluations (best for evaluating teaching and common interest activities) must be distinguished while making sure that local evaluation structures include a sufficiently high percentage of outside experts.

Individual evaluation of full time and academic researchers in theoretical and experimental sciences

Summary and Recommendations

Evaluation of research units and researchers has been practised for a long time already and is now considered a normal process by the scientific community. Its expanding use has become a topical issue due to the recent French law, the Universities' Freedom and Responsibilities (LRU) law that transfers new powers to the universities at the local level. The French university context is complex, work conditions for academic researchers are far from homogenous, yet they have to be evaluated nationally despite differing infrastructures, equipment availability, student educational background and course options. With regard to this specific context, the Académie des sciences has put forward some recommendations based on three important principles: competence, transparency and ethics.

1. The code of ethics

- Evaluators' mandate should be short (3 years) with a renewal on a yearly basis of one third of all committee members.
- Committees should include one expert from outside the field and a high proportion of examiners from other French or foreign institutions (the LRU law specifies 50% in the case of recruitments).
- The procedure and criteria used in an individual evaluation, as well as their adaptations to specific fields or sub-fields, should be posted at the national, institutional and university levels.
- A special effort should be made to identify conflicts and common interests that are not immediately clear, and any ethical issue should be brought to light in advance.
- Each member of an evaluation committee has an obligation of confidentiality; the president is the only person authorized to give more detailed information in case of dispute.
- The full report should be communicated to the examinees without any modifications, confidentiality as to the report's authors being preserved by the evaluation committee and its president.
- A commitment to the code of ethics should be signed by each evaluator. Any failure to comply to the code of ethics should be considered serious professional misconduct.

2. Criteria and tools for evaluating research activities

Any evaluation of research quality and productivity should integrate several levels of analysis.

Qualitative evaluation

Qualitative analysis is the most important facet of an in-depth evaluation. It should be based on an analysis of the scientific work and if necessary augmented by a timely interview. Bibliometric data and other quantitative criteria cannot be a substitute for an evaluation by peers, however once the data has been fully examined and understood, it can help decision-making.

Quantitative evaluation

Bibliometric indicators may be quite useful if used properly, readjusted to the context of the field and integrated into a qualitative evaluation.

- Bibliometric indicators should not be used alone to establish a ranking.
- Greater importance should be given to article citations than to the impact factor of the journal in which the work is published (except in the case of young researchers). The *h* and *g* indices based on citations are useful but of limited interest and should be complemented with new indicators.
- The number of authors in a citation should be taken into account as well as the place of the author's name in fields where the order is not alphabetic.
- The Académie des Sciences suggests organising an inter-organism and interdisciplinary action, together with the science and techniques observatory (OST) and the national agency for higher education and research evaluation (AERES), to reflect on the use of bibliometric tools and the creation of new indicators. Tools currently used should be validated with retrospective tests.

Other criteria of recognition

The scientific quality of a researcher can be evaluated based on the basis of criteria other than those relying on bibliometry, in particular managerial, supervision and leadership skills, the writing of academic books or books for the greater public, the number of languages they are translated into, invited talks to conference plenary sessions, leading participation in international programmes, presidency of an international scientific association, chief-editor positions in international journals, award of significant contracts, awards of prizes and national or international distinctions, membership of French and foreign academies, other distinctions such as nomination to the Institut Universitaire de France, organization of summer schools, symposia, high level international meetings, etc.

Evaluation of research applications

In contrast to what is observed in other countries, and although much progress has been made in this respect, industrial projects and applications are not sufficiently taken into account when evaluating researchers in France.

- Industrial application should become an essential evaluation criterion for those involved in applied research and it should become a factor leading to promotion to the same extent as publications.
- An evaluation scale should be established giving a significant place to the relevance of the research.
- Criteria for evaluating research outcomes that do not directly lead to immediate applications, such as software and prototypes, but that are nonetheless important should be defined.

In the end, it is peer committees that review the evaluation criteria mentioned above, and they should do so mainly based on a personal analysis of publications and interviews. Such an evaluation should include quantitative indicators but also take into account the novelty of the research and its relevance.

3. Criteria for evaluating teaching activities

The LRU law and the recent decree of 23 April 2009 that defines the regulatory measures applicable to academic researchers establish the obligation of evaluating three types of activities: research, teaching and common interest activities. This is made necessary by the fact that the relative importance of these three types of activities may vary during a career. Concerning the evaluation of teaching in all its forms, the Académie des Sciences recommends the following:

- Evaluation of teaching activities may be carried out following several approaches that lead to the production of an evaluation scale at the local and national levels, the local evaluation being the most relevant. An important criterion is student rating of courses, a delicate point that may lead to perverse effects.
- The evaluation of teaching activities should also include objective criteria such as content and novelty of courses (publication of teaching material, manuals, posting of courses and lab material on-line, exhibits, etc.)
- An annual record of teaching obligations should be published each year by each institution, and teaching exemptions should be clearly mentioned and justified.
- The institution should publish each year the percentage of students that successfully finished their study requirements (L1, L2, L3, M1, M2, doctorate) and what they moved on to do, as well as their possible employment prospects per level of studies and at the end of

the thesis. These elements should be taken into account as much as possible to evaluate academic researchers.
- The best *maîtres de conférences* (associate professors) who devote most of their time to teaching (initial training and continuous training) and are unanimously recognized for their pedagogic qualities, should benefit from local promotions such as *Hors Classe* (Exceptional Teacher) or receive bonuses from their institution. Inversely, those who neglect their teaching duties should bear the consequences.

4. Evaluation of common interest activities

- Administrative and common interest activities should be taken into account when evaluating academic and full-time researchers under the new regulation in force, in particular regarding activities that require responsibility (coordinating the first academic year, department leader, international cooperation missions, advising students on courses and jobs, cooperation with industry, patents, promotion of scientific and technical knowledge, etc.).
- Institutions should publish a record of non-teaching responsibilities fulfilled by academic researchers.
- Currently, there are no objective criteria to evaluate these activities. A specific scale should be established to evaluate common interest activities.

5. Evaluation frequency and format

The current frequency of evaluation is too high. The Académie puts forward the following recommendations:
- In-depth evaluations and routine performance assessments should be distinguished.
- The number and frequency of in-depth valuations should be limited to the important steps in a researcher's or teacher's career, *i.e.* recruitment and important promotions and transfers .
- Recruitment is a key step because the staff recruited will become a *fonctionnaire d'État* (civil servant).

  *A nation-wide two-step process should be adopted, with a first cut-off on dossier and a second cut-off after an oral presentation followed by an in-depth interview. Because scientific creativity and novelty are hard to judge based solely on bibliometric data or prepared presentations, the in-depth examination by peers should take on this evaluation role. The ability to teach should also be tested based on pertinent seminars.
- The lists of national qualifications created to make up for the heterogeneity of thesis levels and habilitations to direct research (HDR) are not fully satisfactory. With the recently instituted autonomy of universities, a significant redefinition of the criteria required for obtaining these qualifications should be undertaken. The essential role bestowed onto

graduate schools and university scientific committees should also be redefined. These recommendations must rapidly lead to diplomas recognized for their quality.

*Due to compulsory preliminary registration on aptitude lists, there are four evaluation steps involved in becoming a professor in France, as compared to only two in other countries similar to ours. A general reflection on this topic should be carried out.

- Performance evaluations should be limited and done as part of the standard four-year university activity contract, on the basis of the simplified form used to monitor the normal activity of the staff.

6. Evaluators

- The scientific competence of evaluators is fundamental. The Académie des sciences proposes that a list of prerequisites be prepared and published by the AERES for each category of evaluators.
- Important measures should be taken to ensure the good will of the best evaluators, by making their task easier, reducing the duration of the mandates and making sure this activity is taken into account when appraising administrative or common interest responsibilities.
- All evaluators should be evaluated to guarantee their competence.
- Although evaluation committees specific to each university should be managed locally, they should include a significant number of external examiners (clause specified in the LRU law in case of a recruitment panel)
- The respective roles of national evaluations (essential for research activity) and local evaluations (more appropriate for evaluating teaching and common interest activities) should be distinguished.

7. Follow-up on evaluations

- One of the major difficulties of the evaluation system is that there is frequently absence of an impact. Hence, care should be taken to only perform evaluations when these can lead to a promotion or a career reorientation.
- A distinction should be made between assessing the quality of an activity and the progress of a career and avoid mixing up evaluation and reorientation.
- Careers should be monitored by *ad hoc* committees that take on the role of career counsellors. This system for human resources management should be adopted by each university.
- Universities should post their scientific and educational specificities so that calls for candidates are unbiased.

Individual evaluation of full time and academic researchers in theoretical and experimental sciences

Preliminary remarks

As mentioned in the title, this report only concerns the fields of theoretical and experimental science, from mathematics to biology, medical and pharmaceutical sciences. The report does not take into account humanities and social sciences, although they share with other sciences similar evaluation problems, but also have evaluation issues specific to their fields.
The report concerns the evaluation of both full-time researchers (*chercheurs,* C) who are employed by national research institutions and academic researchers (*enseignants-chercheurs,* E-C) who are employed by the Ministry of Higher Education and Research and normally have both research and teaching duties. Throughout this report, the term “establishment” should be understood as referring to universities while “institution” refers to research organisms. The report also concerns the teaching establishments and research institutions that depend on ministries other than the Ministry of Higher Education and Research.
Finally, we would like to recall that evaluation of groups is outside the scope of this report although the complex interdependence between the activity of a C or a E-C and that of the group or laboratory he/she is affiliated with should not be underestimated.

Introduction: critical presentation of the situation

Individual evaluation is an essential step by which it is possible to assess the competence of full-time and academic researchers and the value of the results of their research activity as well as base recruitment, promotion and performance throughout their career on objective criteria. In contrast to the situation in other countries, there is no real consensus in France as to the evaluation process in spite of the fact, which in itself constitutes significant progress, that evaluation is nowadays accepted by the whole teaching and scientific community. This problem has become particularly topical since the law on the autonomy of universities (LRU law) came into force conferring new powers to universities on a local scale. Several aspects should be considered.

First, an evaluation can only be accepted if it is based on three fundamental principles, namely *competence*, *transparency* and *ethics*, and if it is carried out under by peers.

The *competence* of the evaluators is an unavoidable prerequisite to a good evaluation process and should be guaranteed at the local and national levels, which is not always the case. Similarly, *transparency* of the evaluations, in particular concerning the evaluators’ report, is not always guaranteed. Finally, the standards used, in particular the rules of good practice and ethics, are not always clearly posted. This failure to abide by a code of ethics may lead to conflicts or to biases in favour of certain individuals or groups and can lead to a bad image of the evaluation process.

It is a peculiarity of the current French system that too many rather than too few evaluations are carried out; these take place many times during a career, for instance at recruitment, enrolment on qualification lists, transfers, promotions, when applying for grants. This high frequency of evaluation leads to a huge waste of time and energy and it diminishes the impact of the evaluation itself.

The introduction of quantitative criteria based on bibliometry into the evaluation process of scientific activity constitutes a progress. However, it may lead to important biases due to inadequate use of the data and excessive reliance on this data as compared to other evaluation criteria. Evaluation and bibliometric measures have often been confused. It is necessary to recall that quantitative criteria only become meaningful through the final evaluation of peers.

Except for recruitment, transfers and promotions, personal evaluations are rarely acted upon, which leads to a certain weariness on the part of experts called upon and to their frequent refusal to participate, especially those that are the busiest and most competent.

One issue will be further added and stressed throughout this report, the fact that the context of evaluations is very different according to the field.

In the case of teaching and common interest activities, the situation is even more complex due to the lack of objective evaluation criteria.

All these difficulties are made worse by complex and expensive administrative procedures, to which should be added the time spent by evaluators and evaluatees, all for a process of limited efficiency and often disputable results.

Finally, full-time and academic researchers employed by universities and public research institutions are *fonctionnaires d'État* (civil servants) which creates complications in the follow up to a personal evaluation, in particular for promotions, specific bonuses and teaching exemptions, both at the local (essentially universities) and national levels (national university council (CNU), research institutions, national commission for Enseignants-Chercheurs of the Ministry of agriculture (CNECA) and other national committees).

Furthermore, the university context is complex, work conditions for academic researchers are far from homogenous, yet they have to be evaluated despite differing infrastructures, equipment availability, student educational background and course options. These conditions are made even more difficult by the fact that the teaching load is heavy and that a large number of academic researchers do not belong to institutionally recognised research units.

Considering all these elements, within the current legal framework, in particular the framework set by the LRU law and its implementation decree of 23 April 2009 concerning academic researchers, the Académie des Sciences has deemed it important to propose a number of recommendations based on the best practices currently used by other countries, in the hope that these will contribute to the creation of a European evaluation area.

These recommendations concern successively various facets of a personal evaluation, namely the code of ethics, research, teaching, common interest activities, evaluation frequency and format and finally the follow-up.

1. The code of ethics

One of the main difficulties of the evaluation process is to avoid as much as possible any conflict of interest, however imperceptible and indirect. All countries comparable to France have defined a code of good practice and ethics to counterbalance these potential manipulations of the procedure. France lags behind on this point which is so crucial to the credibility and acceptance of an evaluation procedure by the scientific community. AERES has established a set of good practice and ethics rules, however these were devised to evaluate research institutions, universities, laboratories and research groups, not individual researchers. Currently, there is no set of rules for personal evaluations.

Concerning more specifically good practices and the code of ethics, several important points should be mentioned:

- The evaluators' mandate should be short (no more than 3 years) so as not to penalize the evaluator by the rules of the code of ethics which impose that he/she cannot be both judge and participant. Thus, throughout the mandate, an evaluator cannot apply for a promotion, participate in the recruitment of a new colleague or the promotion of member of his/her group. In order to avoid a conflict of interest, it is not enough that the person concerned leaves the debating room when his/her case is examined because it is well known that he/she will weigh on the final decision. A further reason for limiting the duration of the mandates is that evaluators are required to examine dossiers in-depth, which is difficult to impose over the long term. The idea to renew the members of evaluation committees by thirds every year would allow both to preserve a memory of previous evaluations and to break up lobbies.

- The input from someone outside the field is also advisable (for instance, having a physicist on a biology committee). It is also necessary to have a significant percentage of experts from other universities, 50% in case of recruitments as stipulated by Article 25 of the LRU law. The outside experts should include a percentage, to be defined according to the context, of foreign experts whose opinion may be obtained by remote consultation or in writing. The opinion of the foreign experts will be most useful if based on the comparative appraisal of several candidates. Furthermore, reference letters requested by the committee from the referees suggested by the candidates are very useful. These procedures must apply to all universities, whatever their size, in order to avoid inside recruitment or unjustified local promotions of their own staff. The procedures mentioned above concern universities but could be extended also to France’s main research institutions.

- To ensure transparency, the procedure and the criteria used for personal evaluation should be published, along with their specific modifications for each field or sub-field at the national level (branches of the national council of universities (CNU) and the national committee for scientific research (Comité national de la recherche scientifique) and local level (mainly universities).

- A specific effort should be made to identify conflicts or community of interests that are not immediately apparent. Currently, a researcher may apply for a grant from the national research agency (ANR) for a specific programme and be asked to evaluate grant applications for the same programme. This is a typical situation of conflict of interests because of the competitive nature of the process. If one of the dossiers to be examined poses an ethical problem to an evaluator, he/she must declare so to the president of the jury before the beginning of the evaluation. The latter will decide on the procedure to follow and on the necessity to replace the evaluator.

- All evaluation committee members should respect the clause of confidentiality as to the content of the panel’s debates. In the case of a dispute on the outcome by the applicant, only the president should be authorized to give further details if necessary, while preserving the anonymity of the members of the committee. This will allow evaluators to express their opinion more freely in the report since the anonymity of the author will be preserved and the assessment will be made in the name of the committee as a whole. This procedure would also allow greater transparency. The Académie des sciences considers important that full reports be given to the examinees without any modifications (except for personal remarks). The Académie des Sciences is against posting these reports on-line on an open site.

- Finally, a commitment to abide by standards of good practice and a code of ethics should be signed by each evaluator as is done regularly in other countries. A code of ethics has already been established (in particular by the European Research Council) and can serve as a model. Some rules have been briefly stated in recent French decrees concerning the national council of universities (CNU); however these were not described with any precision, no reference was made to an obligation to sign them, unlike other countries, and nothing was mentioned in case of non compliance. Any breach of the standards of good practice and the code of ethics should be considered serious professional misconduct.

2. Criteria and tools for evaluating research activities

2. A – Introduction: the primordial importance of qualitative evaluation

Important changes were recently introduced in scientific evaluation. Quantitative criteria based on bibliometry, that is the use of scientific publications as indicators of an author's productivity, can now be used; the importance of applications derived from research is more widely recognized; we have learnt to take better account of marks of recognition by the international community. Nonetheless, a qualitative evaluation based on the direct analysis of the work of a researcher, which some have wrongly thought could be partially replaced by bibliometry, should remain an essential step of the evaluation process. In this report, we will nevertheless discuss quantitative evaluation at length in order to remove the ambiguity that frequently exists between bibliometry and evaluation.

The various points that will be discussed successively concern both full-time and academic researchers[1].

2. B – Quantitative evaluation of research: bibliometry

Following the creation of the *Science Citation Index*, the first large bibliographic database, in the 1960s (Annex A), Eugene Garfield had the idea to use article citations as a criterion to define scientific fields[2]. Later, the use of citations led to the construction of tools to measure the relative importance of scientific journals (impact factor). More recently, the use of citations was extended to individual scientific evaluation in particular using the *Web of Science* database commercialised by Thomson Reuters and *Scopus* commercialised by Elsevier. Reliance on bibliometry grew rapidly. However, one should keep in mind that bibliometric tools are still rather rudimentary.

When discussing the quantitative measurements used to evaluate scientific works and their authors, a distinction should be made between evaluation of the sources (*i.e.* the journals) and evaluation of the authors.

- The first is based on the importance of specific journals, identified as the impact factor *IF* published by the database *Journal Citation Report (JCR)* which is a subset of the *ISI Web of*

[1] AERES has proposed a definition useful for an academic researcher based on a minimum number of publications, which constitutes an interesting but insufficiently discriminating criterion, the baseline chosen being set too low. The value is relative and varies according to the field. Furthermore, remarkable researchers may publish very little or not at all for a period of time before communicating a major finding, in particular in mathematical sciences. Finally, some researchers delay publications so as not to hinder a patent application.

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

Knowledge database (see Annex B). Such a practice is dangerous because even the most prestigious journals include mediocre articles that are seldom cited. An article does not become important by the simple fact that it is published in a high profile journal. The above criticism concerning the value of the impact factor in the context of a personal evaluation should not discourage researchers and E-Cs from trying to publish results they consider novel in reputable journals. It is important that French scientists publish as much as possible in the best journals, which are still, in spite of the above comments, an international show-case for many scientific disciplines.

- The second is based on measures of article citations. It has developed very rapidly since since 2005, following the proposition by J.E. Hirsch to associate an indicator *h* with each scientist. Based on the number of most cited articles, this indicator can be calculated very easily with the software provided by the *Web of Science* database of Thomson ISI (Annex C). Another slightly more complex indicator that gives greater importance to excellence, the *g* indicator, was proposed by L. Eghhe in 2006 (Annex D).

What use should be made of citation measures for a personal evaluation of research? Bibliometric indicators can be very useful when they are well used, *i.e.* when they are considered within the context of a discipline, taking into account the size of the particular scientific community and included as part of a qualitative evaluation (see below). Bibliometry can be a relevant tool if it is used with care and by scientific experts. In any case, because of its wide current and projected use, it should not be dismissed altogether which would be an excessive step but its limitations should be stressed, its use and transparency should be improved and new indicators proposed.

Bibliometric indicators should not be the sole basis of a ranking process, because of their intrinsic weaknesses linked to the heterogeneity of individual situations and because of the importance of qualitative evaluation discussed below. Currently, bibliometry suffers from several weaknesses:

- Bibliometric indices differ widely according to the field. They are used rarely in mathematical sciences and immoderately in biological sciences. Furthermore, the mean value of the number of citations, which determines the impact factor, and of the indicators mentioned above is not the same in each field. Criteria must therefore be defined and published for each of the fields. To optimize the use of bibliometrics in the evaluation of a researcher, he (she) should not be associated with a figure but rather be assigned a position indicative of the researcher’s place among fellow researchers in the same field.

- The indicators may influence researchers’ behaviour as some of them might chose to steer their activity in such a way as to get articles accepted in high impact factor journals rather than engage in original and creative research (see Annex E).

- The number and the order of authors on a paper vary according to the discipline (authors may be listed by alphabetical order, seniority or other criteria). This point may be a serious problem. For every citation, the number of authors should be taken into account and in disciplines where the order is not alphabetical, the position of an author's name should also be considered.

- Errors are often made when computing the indices, and this is particularly serious when it is done without the knowledge of the person concerned. A structure similar to the Sciences and Techniques Observatory (OST) should be created to validate the indicators calculated.

- Finally, bibliometric indicators should be used with great care in the case of young researchers. This tool should not be used in the first 5 or 10 years of professional activity depending on the discipline.

In conclusion, indicators cannot substitute for an evaluation by peers but they can, once they have been well scrutinized and understood, become one of the elements to help decision-making. Abnormal values (such as excellent researchers who publish a small number of first-rate articles or inversely researchers who are authors of highly cited papers to which they made only a small contribution) should not disqualify bibliometry as a whole; these are specific cases revealed by bibliometric indicators that justify further analysis by peer committees. The Académie des Sciences is for a reasoned use of bibliometry in a personal evaluation process only if it is part of a peer evaluation and under the control of evaluation committees. It recommends also to develop a series of new indicators in order to adapt bibliometrics to the variety of situations presented by different fields and to follow the evolution in time of this set of indicators when assessing the development of research careers. In order to counterbalance the weaknesses mentioned above, the Académie des Sciences suggests to launch an interdisciplinary and inter-organisational consultation with OST and AERES to examine the use of bibliometric tools and the creation of new indicators. It is also important that the new tools be validated with retrospective tests.

2. C – Other criteria for evaluating research

The performance of a researcher may be evaluated using many criteria other than those based on bibliometry, as witnessed by the remarkable application form used by the European Research Council (ERC) for individual grants, and as recommended previously by the Académie des sciences in its report on attractiveness of research careers « *l'Attractivité des carrières de recherche* ». The new decree of 23 April 2009 implementing the provisions applicable to academic researchers stipulates that from now on all the academic researchers' activities should be taken into account. With respect to the evaluation of research activity, the following main criteria have been listed:
- ability to organize, coordinate, lead research (number and quality of theses supervised, quality and career of post-doctoral students, …)

- writing of scientific books
- writing of books for a wide public,
- contribution of chapters to collective books,
- translation of these books in other countries,
- presidency of international programmes or international scientific associations,
- teaching in summer schools,
- invitations to speak in plenary sessions at meetings,
- position as Chief Editor for international journals,
- awarding of contracts, especially international contracts,
- prizes and national or international distinctions,
- election to a French or foreign academy, to the Institut Universitaire de France, etc.,
- organisation of high level or innovative international symposia and meetings,
- participation in committees or national and international agencies,
- actions concerning the general public.

The particular case of research supervision deserves a specific comment because it has several facets: a facet concerning research directions and strategy, a purely administrative facet which can be heavy, and finally a more complex facet concerning human resources management (see also § 7 on the use of research). The careers of students trained in the laboratory (for instance, at a period 3 years after their thesis) should also be taken into consideration in an individual evaluation when examining research supervision.

All activities that contribute to the smooth progress of research should be recognized by the scientific community. Duties other then research should also come under consideration for the evaluation of a researcher's career. All these criteria need to be interpreted with care and adapted to each person's situation.

2. D – Criteria for evaluating applied research

In countries with strong economic activity, research results and applications are increasingly important due to their direct impact on high technology processes and products. France is somewhat behind in this respect; one way to reduce the gap would be to take into account research applications when evaluating full-time and academic researchers for promotion.

1° Divergent attitudes in matters of applied research

In contrast to other countries, in France, fundamental research is looked upon more favourably than applied research. The creation of prototypes and complex software is valued very little and the involvement of a researcher in developing applications is not considered a natural extension of research. This produces a lag between fundamental knowledge and the applications that could be derived from it. When applications are attempted, they often do not come to anything because of a lack of solid logistic help, interested industrial partners, adequate financing and perseverance. If bibliometrics were to become an essential evaluation criterion, researchers might be led to consider publishing valuable applied work in spite of the fact that doing so might prevent a patent application.

However, the situation is evolving and important progress has been made due to the law on innovation as well as the creation of financing agencies and greater facilities for researchers to take a leave of absence or a secondment to create a business. However, the progress made is not sufficient. The reasons are persisting procedural complexity, the high number of persons involved and the reluctance of research institutions to establish partnerships with industry in a way that answers to the economic logic of business and private investors. Another reason is that products and industrial or societal applications are not taken enough into account when evaluating researchers' careers. It is therefore absolutely necessary in the case of researchers who devote their time to research applications, innovation and other scientific developments, that these become essential evaluation criterion.

2° Evaluation criteria

In the area of *research applications*, the main issue is that of scientific or industrial utility. To define it, researchers could be helped by a scale based on the following criteria:

- Patent applications. Although not everything is patentable, the number of patents is an essential evaluation criterion. It is insufficient though. The number of licences and of patent exploitation on the long term should also be taken into account.
- Partnerships with industry and industrial transfer of prototypes.
- Consulting.
- Creation of start-ups, companies and consortiums of related businesses.
- Partnerships between public or private institutions with scientific and technical research public establishments (EPST), the industry and health sector, etc.
- Personal involvement of researchers in these applications.

Research outcomes, i.e. new scientific contributions that have no immediate industrial application are just as important but harder to evaluate. The number of downloads of a free software, often used as an indicator is a weak one. The Académie des Sciences recommends that evaluation criteria be defined for these activities and that the contribution of researchers who undertake high-level experiments, produce new software and prototypes or work on new societal applications be rewarded.

Evaluation is an unavoidable process in the career of both full-time and academic researchers. It can occur in many forms and at various stages, such as promotions, awards, bonuses and other financial rewards. Evaluation should be used as a tool to introduce in our country a true culture of application of research highlighting the continuum that should exist between fundamental and applied research, which mutually enrich one another.

2. E – Peer evaluation

As already mentioned, evaluation should not be limited to quantitative evaluation, whether based on bibliometry, reputation or any other criterion evoked above. A peers' committee should examine the research work in detail, based on direct analysis of the content of published articles or on interviews, as long as the latter are done in a timely fashion. Concerning articles, researchers being evaluated should choose a limited number of their publications that they wish to submit for evaluation (between five and ten). Evaluators, on the other hand, should sign a commitment to read at least two articles among the ones submitted by the candidate as is current practice in other countries. In some cases, in particular when foreign experts are involved, the evaluation may be carried out through the mail or remote communication. However, it would be better whenever possible that it be done within the framework of a committee session, which has the advantage to allow debate and comparison between various researchers working in the same field. The evaluation should include quantitative indicators and it should also closely consider the novelty of the research and its relevance.

Peer evaluation should be based on a dossier presenting the work of the researcher, with a number of preset forms where the researcher lists his/her best publications and presents his/her career and projects, as is currently done in the case of applications to ERC individual grants. Of course, such "self-evaluation" does not require that the researcher give an opinion on his/her career. It might be useful that the dossier, at least in some cases, includes reference letters from recognized scientists, as is currently and efficiently done at the Institut Universitaire de France. Finally, one should keep in mind that any individual evaluation should take into account the scientific environment in which the researcher works.

3. Evaluation criteria for teaching activities

Until now, academic researchers were evaluated only on their research activity, both at the recruitment stage and for promotions. Obviously, the two other main activities of academic researchers should be taken into account also, that is teaching and administrative involvement, as is done in several other countries comparable to France.

3. A – The academic researcher as part of a teaching team

Evaluating the activity of academic researchers is complicated because such activity is developed in close collaboration with the faculty members of the same university, making individual evaluation difficult in some cases. Evaluation is also complicated by the current positive trend to diversify teaching methods and give greater importance to direct interaction with students, including tutorials. These two points, which are widely mentioned in the decree of 23 April 2009, seem sufficiently important to recall their terms before discussing any further the evaluation process for academic researchers.

The Académie des Sciences had already mentioned in a previous report on career attractivity (*Attractivité des Carrières de Recherche* - 2008) how heavy the teaching load is and does not allow, especially in the case of young *Maîtres de Conférences* (assistant professors) to initiate and develop a significant research project, whether the *maître de conférence* belongs or not to a recognised research group[3]. A thorough examination of the situation should be undertaken and decisions should be reached in each university. This does not mean that the teaching load of academic researchers who are not or very little involved in research should be increased. The implementation of teaching units (UFRs) in each university, by year and by discipline, is essential and represents a solution to this difficulty. Each unit should coordinate courses, define their level and teaching objectives and foster exchanges as a way to stimulate the group.

The training of academic researchers is an important point. It is best done by the academic unit joined by the young lecturer. Training, especially in the case of inexperienced young lecturers, could be envisaged under the form of one- or two-year tutoring by an experienced professor who is well recognized by his/her colleagues. Previous experience with teaching assistants and instructors did not always offer good teaching training conditions; teaching assistants often taught without having received any instruction or support by a professorial group, only and sometimes in the best of cases the help of a qualified supervisor. It should be recalled that CIES centres were created two decades ago to give some basic training to instructors. This is an important point to underline because of the significant number of instructor positions attached to research grants. Supervision of the young instructor by a tutor is beneficial both for the instructor and the students.

3. B – Pedagogical experiments

Teaching experiments carried out in some universities have shown that there are benefits to associating researchers with certain teaching activities. Researchers often appreciate having an access to students, which is easy to do in a teaching situation. By volunteering to teach, researchers discharge some of the academic researchers' heavy teaching load[4] while improving their own general knowledge of the field, a knowledge that is necessary to pursue meaningful research. An ideal scenario, experimented with success in some universities,

[3] There are several categories of university teachers, such as for example the professeurs agrégés (PRAG) and Maîtres de Conférences (MdC). The PRAGs' teaching load is twice that of the MdCs (192 hours/year x 2) and they are not required to do any research. The MdCs have a heavy teaching load (192 hours/year) but must be involved in research. There are some rare instances where, in the same university, some MdCs (or researchers) are not or are no longer engaged in research and have only teaching to do in the same field as a PRAG. The French system is paradoxical in that different categories of staff may be doing the same type of work!

[4] Research institutions and universities should make sure that, in this case, the teaching discharge does not lead to a suppression of the bonus for doctoral student supervision, which is the only reward some academic researchers get for being involved in research.

would be to trust the full responsibility of a course (last undergraduate year L3 or first year of a Master's degree levels M1) to a research group (for example, a recognised university research group (UMR) in partnership with the academic team in charge of the instruction level and under the responsibility of the teaching unit (UFR) or of the relevant academic department. Other schemes are possible and should be experimented with.

3. C - Evaluation of teaching activities

The evaluation of teaching activities may be done following various approaches which all lead to the production of an evaluation scale at the national or local level, the latter level being the most relevant. An important criterion is student evaluation in spite of the fact that it may be biased. Students may appreciate some teachers in spite of the fact that their course content is weak, a professor may deliver very clear lectures although the data and concepts presented may be obsolete, all these are situations that students may not be able to judge clearly. An examination of such an evaluation process for all French universities is urgently required and it should distinguish the various instruction levels (undergraduate or Master level) and the format of the teaching (courses, problem classes, laboratory practicals, remote teaching, tutorials[5]). It should be stressed that how good a teacher can generally only be appreciated by his/her students and close colleagues. Therefore, a teaching evaluation scale should be developed in collaboration with students and it should take into account the points discussed above. The conclusions should be taken into account by the faculty (after examinations, end of year meetings, etc.) in order to improve and/or modify the curriculum and limit as much as possible the biases of such an evaluation. Student evaluation of professors is widely practised in Anglo-Saxon countries and in some French establishments such as the Grandes Ecoles, with much profit and benefit both to students and to teachers[6]. The evaluation procedure could be based for example on that of the Delft and Louvain universities and the Swiss Federal Institutes of Technology in Zurich and Lausanne. The faculty should make clear local difficulties linked to the context of the specific teaching environment (student heterogeneity, overfilled amphitheatre and other local material conditions).

The best lecturers (Maîtres de Conférences), who are devoted teachers and unanimously recognized for their teaching qualities, should have the opportunity to apply for local promotions, especially to the *Hors classe* (exceptional class) distinction, as specified in Article 22 of the decree of 23 April 2009 and as is practiced in other countries. Exceptionally good teachers should be rewarded with a special bonus, a kind of teaching award at the level of each university. This would create real competition among teachers to improve teaching procedures. Basic courses, such as first year courses, should also be rewarded as well as the most technical ones (IUT, professional training). Conversely, academic researchers (MdCs or

[5] An incentive to supervise and help students individually should be set up by the teaching unit under a format yet to be defined (tutorials, office hours, etc.)
[6] Pedagogical support by specialized staff should be available to teachers in case of difficulties.

PRAGs) who deliberately neglect their teaching responsibilities (for instance, frequent absences or questionable course quality) should suffer some consequences. From this point of view, the opinion of the students should be communicated to the local faculty and taken into account when examining promotions at the national level. Before this can happen and as mentioned above, evaluation criteria for all levels of teaching, whether university degrees or continuous training, should be made more objective than those currently used and should be specifically adapted for each discipline, if necessary.

Evaluation of teaching activities must include a number of objective criteria such as course content and novelty (pedagogical publications, editing of course manuals, on-line availability of the professor's practical courses and problem classes, foreign distribution or translation of course manuals, exhibits, etc.). Such an evaluation could be based on existing pedagogic archives which should be more systematically updated in terms of manuals and pedagogic material produced by teachers.

To inform evaluators (national and local experts), it is officially required that precise records of the annual teaching load be published each year by every establishment, with a clear statement and justification for any teaching discharge.

Finally, the percentage of successful students, what the students went on to do at the end of each study year, by level and by discipline (for instance, at the end of each undergraduate or Master's degree year and after they completed their doctoral studies) and the employment found after their thesis, by discipline, should be made public and taken into account in an individual evaluation of E-Cs even if it is not always easy to identify the impact of each teacher.

4. Evaluation of common interest activities

The LRU law that confers autonomy to universities differentiates 6 public service missions for higher education. Apart from the traditional missions, namely the initial and continuous education of students, scientific and technical research, communication of results and technological transfer, the law added 4 missions that the university community must fulfil : career counselling and assistance to find a job, promotion of scientific and technological culture and information, participation in the development of a European higher education and research area and international cooperation. The 4 latter missions concern common interest activities. The recent decree of 23 April 2009 that details the terms of the mandate of academic teachers confirms that all their activities should be taken into account and evaluated, especially administrative and/or common interest activities that require a certain amount of responsibility (coordinator of 1st year students, leadership of a teaching unit, research supervision within the establishment, directorship of a school of doctoral studies, of a group of research laboratories, international relations, etc.). These activities should be organized so as to not take up too much of the academic's time and energy. Currently, attendance at the high number of meetings sterilizes the activity of many researchers and teachers or gives them an excuse not to pursue their main mission, research.

Among the common interest activities that are now evaluated, one should mention positions such as presidency of a university, institution or research unit, involvement in counselling, examination committees, tutorials, student career counselling and help in finding employment, collaboration with public or private companies, development, expertise and coordination of fundamental, applied, pedagogic and technological research and technological transfer. Furthermore, the requirement to better explain the nature of science and its results to society at large has led to various initiatives and developments which should be recognized, as provided for by the decree of 23 April 2009. According to this decree, academic researchers contribute to the dialogue between science and society in particular through the diffusion of scientific and technical culture and information. Finally, the function of evaluator should also be taken into account under the category of common interest activity.

A record of such activities should be regularly published by each establishment and for each academic researcher.

Evaluation of common interest activities has become a legal requirement, however the tools to do so have not been specified. It is a new territory that needs exploring. The evaluation procedures used by other public administrations could be used as a guide.

5. Evaluation frequency and format

In-depth evaluations carried out in the context of steps such as recruitment, promotion, change of category or class should be differentiated from regular performance assessments.

In-depth evaluations should be limited in number and frequency, and only used for key moments of a research or teaching career : recruitment[7], competitive examinations, promotions and class level changes.

The number of regular evaluations should also be limited and should correspond to the institutional contract (contrat d'établissement) i.e. every 4 years. These evaluations should be based on the activity records that are used to monitor the normal activity of the staff. Currently, these records exist and are used by academic researchers who are engaged in

[7] Universities must post their scientific and pedagogic specifications so that calls for candidates are not biased. Candidates who make it through the first cut-off and are invite to come for an interview should have their expenses paid by the University. The percentage of outside candidates successfully recruited and their nationality should be mentioned in the four-year report. Recruitment of in-house candidates has had disastrous consequences in some universities.

research within recognized research units (*e.g.* UMR CNRS). A special procedure is used by the scientific and technical research public establishments (EPST) and the public industrial and commercial establishments (EPIC) for permanent researchers. For example, CNRS researchers are required by their contract to fill annual activity report forms. Their examination by the relevant CNRS divisions independently from the evaluation of laboratories and research groups is a lengthy, useless procedure that takes up too much time and involves too many scientific and administrative staff. An evaluation every four years, during the four-yearly research assessment, is more than sufficient. The purpose of these activity records are to make it easier for a university president or an institution director to identify full-time or academic researchers who are encountering difficulties. In such cases, the establishment's head should designate experts to provide an in-depth report on the basis of which, if necessary, they may prepare an interview (the procedure is described below in § 7).

All evaluations, whatever the procedure, should end with the production of a written document in order to keep a record of the event and its history.

Furthermore, the Académie des Sciences would like to comment and make recommendations concerning the two step-process that currently requires academic researchers to request their registration on an aptitude list before they can apply for a position as a Maître de Conférences or Professor. Aptitude lists were created to counterbalance the heterogeneity of doctoral experiences in terms of quality of the theses and theses supervisors. The recently acquired autonomy of the universities should rapidly lead to a complete revision, at the local level, of the criteria that need to be fulfilled for obtaining a thesis or qualifying as a research supervisor. Doctoral schools and scientific councils play a crucial role that also needs to be redefined. Furthermore, evaluators sitting on the above evaluation committees should receive some kind of reward for their time and responsibility. Their professional records should mention the number of committees they belong to, a criterion that should be weighted by the role they have on such committees which is always difficult to evaluate (see § 6). These recommendations should improve the recognition given to theses and thesis supervisors. The concept of a national diploma, dear to the French egalitarian system, contradicts the concept of autonomous universities. There are two consequences: a) a certain disrepute of theses and thesis supervisors by employers in the private sector, the managers of which are usually not familiar with the content and quality of doctoral work; b) a huge and futile work load for candidates and colleagues in charge of evaluations (national council of universities). Finally, the current complexity of the French system should be stressed. Four evaluation steps are required to be promoted to a professorship (due to the registration on aptitude lists) while the same procedure is done in two steps in countries comparable to ours. The whole process should be rethought.

6. Evaluators

6. A – Choice of evaluators

The quality of the evaluators is crucial to the process. Whatever the level of the evaluation, care should be taken to choose evaluators who are as competent if not more so than the person evaluated. The Académie des Sciences proposes that, for each category of evaluators, prerequisites prepared by the AERES[8] should be published. Each evaluator should provide a short CV and, for example, 10 publications (on fundamental or applied work) published in the last ten years (the number may be adjusted for young recruits).

To guarantee competence and scientific excellence, the evaluators should themselves be evaluated locally by the establishment or at the national level (for instance by the national council of universities (CNU) or the national commission for academic researchers of the Ministry of Agriculture (CNECA). One of the major difficulties is the refusal of the best full-time and academic researchers to participate in evaluation committees by lack of time or interest. The issue of time, which is crucial if the best examiners are to participate, may be partially overcome by limiting the duration of the evaluators' mandate to three years, a measure already suggested above on ethical grounds. Concerning motivation, the issue is more difficult but could be solved by recognizing evaluation activities in career advancement. That might help attract the most qualified on the committees.

The issue of financial compensation for evaluators and at what level has often been discussed. It does not seem necessary to compensate evaluators who are French nationals because their travel expenses are fully covered anyway. Some compensation already exists for examiners participating in CNRS recruitment sessions and some amount of compensation will be provided for CNU committees under decree of 23 April 2009 that modifies the provisions relating to the university councils. The more delicate issue is that of foreign evaluators who often are asked to provide important services to the French research institutions. Compensations for foreign evaluators are certainly useful but should be kept within reasonable amounts. Above all, they should be associated with better conditions when foreigners are invited to participate in evaluation in France, in particular concerning reimbursement of travel expenses.

6. B – The respective roles of national and local evaluations

Regarding the respective roles of national and local evaluations, it is clear that scientific activities should be evaluated at the national level (CNU, research institutions and other national committees) for lack of a European evaluation system that would guarantee a larger and more independent pool of evaluators. There should be a provision by which local

[8] In Germany for instance, evaluators of ANR type projects are elected by young researchers as long as they obtained their doctoral degree more than three years previously.

establishments can give their opinion. Teaching and common interest activities should be evaluated locally.

This dual level evaluation is a problem for university evaluation committees which must be managed locally but include a high percentage of members not belonging to the university (in the case of recruitment, Article 25 of the LRU law requires that 50% of committee members be external to the university), the term « external » covering experts from other French or foreign institutions and universities.

Instructions for evaluators should be clearly stated by each establishment and institution.

7. Functional use of the evaluation

One of the main difficulties of our evaluation system is the frequent lack of a follow-up on the evaluation once a researcher or E-C has been recruited. It is therefore important that evaluations be carried out when they may lead to a promotion, career reorientation or another tangible form of recognition (for instance a bonus).

7. A - Recruitment

Evaluation plays a major role during the recruitment process and errors should be an exception in view of the permanent nature of French research and teaching positions since successful candidates become state employees. This is the reason why universities and research institutions should adopt transparent procedures and rely on a panel of competent evaluators (see § 1 et 6). Position descriptions and details of the application process should be posted well in advance so that candidates (French or foreign) applying from outside of France have sufficient time to complete the procedure. Finally, a two-step process should be widely adopted, with a first cut-off based on the application dossier to select candidates to be invited to give a prepared presentation and submit to an in-depth interview. Creativity and research novelty, two decisive criteria for recruitment, are difficult to appreciate based on a bibliometric analysis or a prepared presentation. The in-depth examination by peers should be able to get around the difficulty. Because of the large number of candidates examined, the typical presentation plus interview carried out by the major French research institutions only lasts from 10 to 30 minutes per candidate and is the main reason there are recruitment errors.

Furthermore, universities should participate towards the expenses of eligible candidates so as not to limit applications by young candidates who have few financial means and to make the recruitment commissions more aware of their responsibilities. Finally, it is important to recall that a significant percentage of external evaluators should participate in a recruitment committee (50% as stipulated by the LRU law) in order to avoid in-house recruitment which is very frequent in some universities.

7. B – Promotion

In the case of promotions to higher grades, change of status (researcher to research supervisor or assistant professor to full professor) or in the case of a transfer, the same transparency, competence and ethical rules should apply (see §1 and § 5) and the same percentage of external evaluators should be involved.

7. C – Career monitoring

Finally, the important issue of regular assessments throughout a teacher's or researcher's career should be addressed. In most cases, the research institution or university communicates the report provided by the evaluation committee to the researcher or teacher evaluated, for information and as judgement on the excellence, relevance and quality of his/her work. Specific recommendations may be added by the senior staff in the research or academic institutions to optimize the work of the person being evaluated. However, the current French system by which the teaching and research staff are government employees does not allow any way out if the person recruited turns out to be ill-adapted to the job. Hence, it is important to make room for improvement throughout the career of the person recruited. Modifications to the professional trajectory and the evolution of expertises should be taken into account, and attention should be paid to the personal wishes of the staff member in order to avoid a resignation or some of the uncomfortable situations for which today there are no satisfying solutions. One possibility would be to allow academic researchers to state on a regular basis (every four years) their wishes, in particular concerning the division of their time between teaching and research. They would then be evaluated according to their chosen direction.

7. D – Procedures for career monitoring

When an evaluation brings to light some problems, an *ad hoc* committee under the stewardship of a scientific personality recognized for his/her human and managerial qualities, should hold an extensive discussion with the examinee to help him/her find the best solution to his/her difficulties (locally in the same university or in other public institutions that depend on other ministries and eventually, if necessary, retraining). Each institution and university should create such committees and keep a list of persons who might be able to play this role on a volunteer basis. Such a system could be very similar to a joint commission and should be set up rapidly in each university.

Currently, there are no professional solutions offered. The French system is paralyzed by a regrettable and disastrous confusion regarding the role of evaluation committees, where evaluation of competence, excellence and individual management of careers are too often mixed. If professional problems arise, the search for a solution is *de facto* left to the head of the research (UFR) or teaching (UMR) staff at a university. In general, few satisfying solutions are offered to the person evaluated. This leads to a true waste of competence and human capital in spite of the fact that teaching and research are based on human resources. It is necessary to get around this shortcoming of the French system and develop an experimental mechanism without necessarily creating an additional administrative structure.

CONCLUSION

Evaluation is an unavoidable process which ensures that full-time and academic researchers fulfil all aspects of their mandate, namely research, teaching and common interest activities, in an up-to-date and efficient manner and with the full knowledge that the relative importance of these activities might vary within a career. Evaluation is a complex and delicate process that requires professionalism, a rigorous ethical environment and the acceptance by all of the "rules of the game".

Evaluation, a major tool used for directing and stimulating research and teaching institutions, evaluation is now finally recognized and accepted. Beyond its essential role as a tool for decision-making bodies, it should be viewed also as being an integral part of the production process of research, teaching and common interest activities. Its end goal should equally be to help full-time and academic researchers to advance towards a high level of excellence in their field and maintain it. This external point of view may offer a rare opportunity to review the direction of an academic's work and eventually help reorienting research topics or his/her career. For all these reasons, evaluation should be considered as a positive process by both full-time and academic researchers.

All the major scientific countries have had to face problems associated with evaluations. The solutions differ according to the culture and position of the staff. We recommend that the French evaluation system be linked to a European system in terms of common procedural rules and exchange of evaluators between different countries.

ANNEX 1
COMPOSITION OF THE WORKING GROUP

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 Salpetrière

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 COURTILLOT
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

ANNEX 2
LIST OF PERSONALITIES CONSULTED FROM 16 MARCH TO 30 JUNE 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

ANNEX 3
GLOSSARY

AERES	Agence d'Évaluation de la Recherche et de l'Enseignement Supérieur (National agency for higher education and research evaluation)
ANR	Agence Nationale de la Recherche (National research agency)
ATER	Attaché Temporaire d'Enseignement et de Recherche (Temporary teaching and research assistant)
CIES	Centre d'Initiation à l'Enseignement Supérieur (University teacher training centre)
CNECA	Commission Nationale des Enseignants-Chercheurs dépendant du ministère de l'Agriculture (National commission for academic researchers of the Ministry of agriculture)
CNRS	Centre National de la Recherche Scientifique (French national centre for scientific research)
CNU	Conseil National des Universités (National council of universities)
EPIC	Établissement Public à caractère Industriel et Commercial (Public industrial and commercial establishment)
EPST	Établissement Public à caractère Scientifique et Technologique (Scientific and technical research public establishments)
ERC	European Research Council
HDR	Habilitation à Diriger des Recherches (Habilitation to direct research)
IF	Impact Factor
INIST	Institut de l'Information Scientifique et Technique (Institute for Scientific and Technical Information)
INRIA	Institut National de Recherche en Informatique et en Automatique (National Institute for Research in Computer Science and Control)

IUF	Institut Universitaire de France (University institute of France)
IUT	Institut Universitaire de Technologie (University institute for technology)
LRU	Loi relative aux Libertés et Responsabilités des Universités (Universities' Freedom and Responsibilities law)
OST	Observatoire des Sciences et des Techniques (Science and techniques observatory)
PRAG	Professeur Agrégé (Secondary school teacher qualified to teach at universities)
UFR	Unité de Formation et de Recherche (University department)
UMR	Unité Mixte de Recherche (Recognised university research group)

ANNEX A
WHERE TO FIND BASIC INFORMATION?

Currently there are two major databases, ISI-Thomson (Reuters) and SCOPUS (Elsevier). Both charge a fee for their use. CNRS researchers have access to ISI through INIST and it would be fair to allow access to this database to the university community. Studies have shown that the ISI database, although more modest in size than SCOPUS, contains all the essential bibliographic information since the early 1920s[9]. The ISI database is continuously improving, in particular by keeping track of conference proceedings and books. It represents a serious base from which information may be analyzed using relevant bibliometric tools. In the mathematics field, the Math Reviews database holds more information than the more general databases mentioned above.

[9] SCOPUS covers a much higher number of journals than ISI (16000 and 9000, respectively). However, ISI has indexed journals back to the beginning of the 20th century. Studies have shown that ISI covers 80 to 100% of citations in biochemistry, molecular biology, chemistry and physics. The coverage is only 40 to 60% for mathematics and engineering sciences. The coverage of conference proceedings and books by the ISI database is increasing.

ANNEX B
INDICATORS TO AVOID

The impact factor (IF) of a journal for a given year N is defined by the ISI database as the ratio of the total number of times articles published in a given journal were cited by the journals indexed by ISI during the years N-1 and N-2 to the total number of articles published in that given journal over those years.
The impact factor is not adapted to an evaluation because it can conceal reality. Multidisciplinary journals such as Nature and Science have an impact factor of 28-30 while the IF is on average of the order of 3-6[10] for specialized journals published by scholarly societies. The impact factor of a journal is basically equivalent to an annual average citation index to all the articles published by a particular journal. However, only a very small number of articles receive a high number of citations[11]. Consequently, the average citation number may, depending on the field, differ widely from the average index for that particular field.

Briefly, the use of the impact factor to evaluate an individual researcher should be banished because it is unfounded. Moreover, an over-reliance on the impact factor of a journal during the personal evaluation of a researcher is pernicious because in the long term it will sentence all the lesser scientific journals to disappear.

[10] These values depend strongly on the field and subfield.

[11] In a typical journal, about 15% of the articles provide half the citations to the journal. For information, in a prestigious journal such as Nature, over the two-year period, 77% of articles are cited a number of times below or even far below the impact factor. The percentage is even 80% for Science.

ANNEX C
PRODUCTIVITY INDICES

J. Hirsch at University of California, San Diego, proposed that the scientific performance of a researcher be measured by the *h* index[12]. This indicator rapidly became quite popular and has now become a classic tool that researchers use without being aware of its limits and shortcomings. The same is true of the administrations who have just started using it.

The *h* index is calculated by classifying an author's publications by decreasing number of citations, C_i, so that a paper ranked $i=h$ received at least *h* citations. This index is a number associated to the given researcher and is calculated independently of the impact of the rest of his/her publications (but it may be expanded to groups of researchers, a laboratory, a discipline, a journal, etc.). It takes into account all the articles cited without giving any particular weight to those that have been frequently cited and that are still cited for historical reasons . The main interest of the *h* index is that it eliminates from the evaluation the numerous articles by that author that are never or little cited. Many normalized variants of the *h* index have been developed that take into account the seniority of the authors or the number of years the researcher has been active such as for example the *m* factor calculated as *h* over the number of years since first publication.

Its easy calculation from bibliographic databases has contributed to the great popularity of this index among researchers. It is even calculated by a one-click procedure using the *Web of Science* software in the ISI-Thomson database. This apparent ease may however be misleading to an ill-informed user. As with all bibliometric indicators, the main difficulty left to the evaluator resides in the relevance of the initial bibliographic data covered. Experts should carry out preliminary work to provide the best choice of scientific fields, suppress homonyms, get around the problem of hyphenated names[13], etc.

The *h* index has many shortcomings. It does not allow to detect excellence when a researcher

---

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

[13] There are several ways to use the ISI Web of Knowledge database. The first (and easiest) consists in using the Web of Science Advanced Search option which gives for a given author the list of articles published in the journals indexed in the base. The second is more delicate and uses the Cited Reference Search option which gives for a given author the list of published articles even if these were published in journals or conference proceedings not indexed by the base. Because non-existence in a base is often due to the negligence of citing authors or typographical errors in the article's references, the first option may lead to a significant reduction in the number of citations.

is highly productive. It is not adapted for instance to researchers who have published a small number of crucial articles that are very frequently cited since the index does not give them any particular advantage. Thus the *h* index cannot distinguish between two differing profiles of researchers with the same *h* value, namely one with a modest number of citations until reaching an article of *h*-rank and the other with a high total number of citations for the *h* first publications. Furthermore, the index does not change when the first *h* articles that received $\geq h$ citations are cited more frequently with time. This is often the case because the most cited publications benefit *de facto* from higher visibility, a situation that further increases their chances of being cited in a kind of positive loop system.

ANNEX D
INDICATOR OF PRODUCTION EXCELLENCE

Some shortcomings of Hirsch's index have been corrected by L. Egghe who proposed in 2006 a *g* index giving greater importance to the quality rather than the quantity of scientific publications[14,15]. Because articles are ranked by the decreasing number of citations C_i, an article (real or virtual) ranked *g* is such that the total number of times it was cited in the *g* first articles is higher or equal to *g* squared[16]. The *g* index which is obtained by an integral of the citations up to the rank *g* takes better account of the impact of highly cited publications, even if their number is limited, than the *h* index which is limited by the total number of papers published. The *g* index emphasizes quality. To this day, it is not provided by the databases and has to be calculated from each bibliometric file[17]. As for the *h* index, preliminary work is required to clean up the bibliographic files before calculating the *g* index.

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

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

[16] Rigorously, the conditions $\sum_{i=1}^{i=g} C_i \geq g^2$ and $\sum_{i=1}^{i=g+1} C_i < (g+1)^2$ are fulfilled.

[17] The website of the Académie des Sciences will soon provide a calculation assistant for these indices.

ANNEX E
BIASES AND MANIPULATIONS

Very early on, researchers have attempted to improve their bibliometric ranking through manipulation of the indices, for instance, by favouring publication in prestigious journals with a high impact factor when this factor is taken into account for their evaluation. Such a behaviour affects the writing and format of the paper raising doubts as to the truthfulness of the information communicated. Some indicators may incite researchers to follow known paths rather than engage in creative research. When citations are taken into account for an evaluation, the case of authors citing one another has become frequent practice as is excessive self-citation and the division of a long paper into several short articles (salami slicing). It is easy to identify such "citation experts" by the fact that their articles often have a lower citation index than the impact factor of the journal in which they are published.

Furthermore, some journal publishers have no scruples in recommending to their authors to cite as much as possible articles published in the same journal so as to increase the impact factor of the journal.

The number and the order (alphabetic, by seniority or other factor) of the authors on a publication vary according to the field. This point may be problematic when an author's contribution to the publication is very small. This is a point that should be carefully examined by peers.

Another important point is whether to normalize citations to take into account the number of co-authors. The answer is far from simple. Normalized indices undoubtedly differ significantly from the original ones especially in fields where publications frequently involve a large number of co-authors or when comparing theoretical and experimental researchers. Would the indices thus modified still be reliable? Advances may come from weighting of an author's index by the contribution made to the work published. For now, as already mentioned, the information content associated to the value of the index can only be appraised by peer examination.

For the reasons discussed above, indices need to be developed in a way so as to overcome or at least minimize biases and possible manipulation.