# Questioning scientific evaluation.

Last recall: Deadline : March 1st 2006. Call for paper (english) — Appel à contributions (français) — EspacesTemps.net lädt ein, Beiträge einzureichen (deutsch).

Par Responsable éditoriale

---

Questioning scientific evaluation.

Evaluation has taken on a major role in scientific production. By default, what one calls ‘evaluation’ mostly refers to the quality assessment of just-produced works — an assessment carried out within an evaluative structure considered as independent of logics exterior to the interests of knowledge. In physics, biology, in technical disciplines and, to a lesser extent, in mathematics, as well as in these of the social sciences that strive to follow the ‘hard sciences model’, a ‘standard’ evaluation system has been set up during the last decades. It appears as a perfectly legible procedure in which every article, submitted to a journal is examined by referees designated by the journal’s editorial board among the specialists of its field of research. It is upon the judgment of these very referees that depends, in a mechanical way, the decision to publish or not to publish. Although very simple in its principle, this system carries heavy consequences.

As a hypothesis, one may say that the rising success of this method has had, as an effect, a hierarchical classification of journals in reference to their legitimacy. As a result, emerged an effect of positive retroaction integrating all of the evaluation devices within a coherent system. Its logics can be interpreted in the following way:

Firstly, an evaluation device, recognized for its competency and equity, defines *ipso facto* the quality of a journal that includes this device in its operating structure. The journals are then themselves evaluated in the same manner by other evaluators. The occurrence of bibliographical references to an article within other articles, measured by the ‘science citation index’, provides thereafter a quantitative means of classification, according to the text’s impact factor, while taking into account the reputation of each of the relevant media. The evaluation of researchers and research teams — and, by extension, of departments,

faculties, and universities — can then be immediately deduced from their publication or citation success within the best classified journals. This can result in an automatic public or private financing of these very actors and reinforce their capacity to produce results destined to be awarded by the evaluation device.

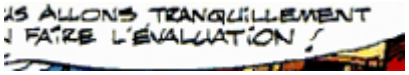
Lastly and above all, by its very success, this device has negative effects on all excluded elements: on refused articles, on journals that do not subscribe to the evaluation norms, on books, on unclassifiable publications and, even beyond, on research activities that do not lead to productions publishable according to these norms… Evaluation, starting with the evaluation of ‘scientific’ journals, appears thus as the organizing force of scientific legitimacy itself of the entire world of scientific research.

This situation is fascinating for researchers in social sciences in at least two ways. Firstly, because they are increasingly concerned on a personal level. Secondly, because it constitutes a fascinating object for their own investigations.

It is in this context that EspacesTemps.*net* launches its call for contribution concerning the following theme:

## A scientific evaluation’

### Can dynamics of knowledge be normed’




This general interrogation is articulated in a series of questions.
1.What and how is being evaluated and who evaluates? Beyond the judging of submitted texts, one can question the evaluation of any production: PhD thesis, science books, large public essays, research projects, research teams and institutions, products not belonging directly to the scientific world (university and scholar manuals, philosophy, architectural or urban projects), extra-research activities of scientists… In this way, one can question the scientific component of any kind of social production.

2. What are the strengths and weaknesses of the ‘standard’ system? Does it contribute, in a positive way, to the scientific independence of scientists and does it take the innovative aspect of their work into account? Do norms and set up hierarchies represent a risk of ignoring the most original among contributions? Do they represent a hindrance for interdisciplinary research? Are there malfunctions within this system that alienate its objectives? Does today’s journal market guarantee for plurality or, on the contrary, does it unduly tend to impose *doxas*, due to its oligopolistic structure? Is there any compatibility between the standard system and a dynamic editorial policy’

3. Can one fix evaluation norms independently from scientific criteria and their dynamics? Or, on the contrary, shout one first define the latter in order to deduce from them different modes of evaluation? Could one — and should one — evaluate questions rather than answers, conjectures rather then results, global theories rather than punctual advances’

4. Is there a specificity of human and social sciences in regard of questions of evaluation and evaluability? If so, what is this specificity and what is its base? Are we moving towards a convergence of the evaluation modalities in all of the research domains and disciplines’

5. Can a system of scientific evaluation be evaluated, by analysing its device or the effects induced on scientific results?

6. Are there alternatives to the standard system? Are there other modes of paper production, (see, for example the ‘Public Library of Science’ project)? Are there other types of evaluators outside of the academic world… deserving attention? Does the emergence of new scientific media, such as the internet, modify the relation between the evaluators and the evaluated? Can the evaluation take into account research components appearing as peripheral but in fact fundamental, such as epistemology, technological consequences, scientific culture or citizenship? Can one imagine other actors than researchers participating in the scientific evaluation process’

7. Must one evaluate science in order to make it innovative? By what criteria can we tell that science is in progress? Are there means, other than evaluation, by

which one can make sure that scientific organisations respect the ‘ethical contract of knowledge’ with society and that they make good use of the means with which this society entrusts them? Does the intervention of exterior actors, such as governments or of citizens, necessary lead to a dangerous social control on research’

8. How can one integrate critical reflexivity on evaluation in evaluation itself? What recommendations can one make so that the evaluation of research work and that of the scientific life does, in fact, maximize innovation and production of knowledge’

9. … And, by the way, are there suggestions to ameliorate the evaluation device of EspacesTemps.net?

You are invited to submit texts, *until March 1st 2006*, for publication in a “ Traverse” of EspacesTemps.*net*. These texts can be personal accounts or descriptive analyses, as well as theoretical statements and recommendations. To begin with, you can from now on send your proposition to Redaction@EspacesTemps.net, in the form of an abstract of 2000 signs. You can also suggest an evaluation mode of the text that you are addressing.

---

L’évaluation en questions.

L’évaluation a pris un rôle majeur dans la production scientifique. Par défaut, ce qu’on appelle « évaluation » désigne souvent l’appréciation des travaux au moment même où ils sont produits, par l’intermédiaire d’un dispositif de jugement réputé indépendant des logiques d’intérêts extérieures à la connaissance. En physique, en biologie, dans les disciplines technologiques et, dans une moindre mesure, en mathématiques et dans les parties des sciences sociales qui cherchent à s’inspirer du « modèle de sciences dures », un système « standard », très simple dans son principe et très puissant dans ses conséquences, s’est mis en place durant les dernières décennies. Ce système, très clair en ce qui concerne les articles des revues, se présente comme une procédure parfaitement lisible. Dans ce contexte, chaque article est examiné par des évaluateurs (*referees*) désignés par la direction de cette revue parmi les

spécialistes du domaine du domaine correspondant. Du jugement des *referees* dépend de manière mécanique la décision de publier ou de ne pas publier.

On peut formuler l'hypothèse que le succès progressivement acquis par cette méthode ont eu pour effet de hiérarchiser les revues selon un principe de légitimité, ce qui a engendré un effet en chaîne doté de fortes rétroactions positives, qui portent sur l'ensemble des dispositifs d'évaluation de toutes les productions scientifiques et les intégrant dans un système cohérent. Voici comment on peut en lire les logiques : d'abord, dans le cadre des publications, un système d'évaluation des textes jugé compétent et équitable définit *ipso facto* la qualité d'une revue qui l'intègre à son fonctionnement. Les revues sont alors elles-mêmes évaluées par d'autres évaluateurs sur cette base. L'existence de références bibliographiques des articles dans d'autres articles, mesurée par le *Science citation index*, fournit ensuite un moyen quantifié de classer les textes selon leur capacité de pénétration dans le marché (*impact factor*), en tenant compte de la notoriété de chacun des supports concernés.

L'évaluation des chercheurs et des équipes ainsi que, par extension, des départements, des facultés et des universités peut alors être déduite de manière immédiate par leur plus ou moins grand succès dans la publication ou la citation par les revues les mieux classées, ce qui peut entraîner un financement public ou privé automatique de ces acteurs et renforcer leur capacité à produire des résultats destinés à être primés par le dispositif d'évaluation.

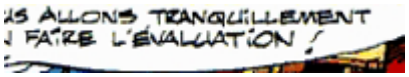
Enfin et surtout, ce dispositif a, du fait même de son succès, des effets négatifs sur tous les éléments qui ne lui appartiennent pas : articles recalés, revues n'appliquant pas ces normes d'évaluation, livres, publications inclassables, et au-delà, activités des chercheurs et des équipes ne donnant pas lieu à des produits publiables selon ces normes, etc. L'évaluation, ainsi organisée à partir de celle des articles de revues, apparaît donc comme la force organisatrice de la légitimité scientifique de l'ensemble du monde de la recherche.

Cette situation est doublement fascinante pour les chercheurs en sciences sociales. D'abord parce qu'elle les concerne de manière croissante, ensuite parce qu'il s'agit d'un objet d'étude fascinant pour leurs propres investigations.

C’est dans ce contexte qu’EspacesTemps.*net* lance un appel à contributions sur le thème :

## Une évaluation scientifique ?

### Peut-on normer la dynamique des savoirs ?




Ce questionnement général s’articule en une série d’interrogations.
1. Qu’évalue-t-on, qui évalue et comment ? Au-delà du jugement porté sur les textes soumis à publication, on peut s’interroger sur l’évaluation du caractère scientifique de l’ensemble des productions : thèses de doctorat, livres scientifiques, essais destinés à un public élargi, projets de recherche, équipes et institutions de recherche, produits n’appartenant pas directement au monde scientifique (philosophie, manuels universitaires ou scolaires ; projets d’architecture ou d’urbanisme), activités des chercheurs en dehors de la recherche… En ce sens peut être interrogée la composante scientifique de toute production sociale.

2. Quelles sont les forces et les faiblesses du système « standard » ? Permet-il une augmentation significative de l’indépendance des chercheurs, de la prise en compte de l’innovation ? Le poids de la norme et des hiérarchies instituées fait-t-il courir le risque d’ignorer les contributions les plus originales et constitue-t-il un frein à l’interdisciplinarité ? Y a-t-il des dysfonctionnements de ce système qui en dénaturent les objectifs ? Le marché actuel des revues est-il le garant du pluralisme ou, inversement, tend-il à imposer indûment des *doxa* en raison de sa structure oligopolistique ? Peut-il y avoir compatibilité entre le système standard et une politique éditoriale dynamique des revues ?

3. Peut-on fixer des normes d’évaluation indépendamment des critères de scientificité et de leur dynamique ? Ou au contraire doit-on d’abord définir ces derniers pour en déduire des modes d’évaluation ? Peut-on évaluer les questions plutôt que les réponses, les conjectures et pas seulement les résultats, les théories globales et pas seulement les avancées ponctuelles ?

4. Y a-t-il une spécificité des sciences humaines et sociales en matière d'évaluation et d'évaluabilité ? Si oui, en quoi consiste-t-elle et sur quoi repose-t-elle ? Va-t-on vers une convergence des modalités d'évaluation dans l'ensemble des disciplines et des domaines de la recherche ?

5. Peut-on évaluer un système d'évaluation scientifique, par des mesures portant sur le dispositif, ou par des analyses des effets induits sur les résultats scientifiques ?

6. Existe-t-il des alternatives au système standard ? D'autres modes de production des revues (voir par exemple le projet *Public Library of Science*), d'autres types d'évaluateurs en dehors du monde universitaire… qui méritent attention ? L'émergence de nouveaux supports de publications, comme l'internet, changent-ils la donne des relations entre évaluateurs et évalués ? Peut-on prendre en compte dans l'évaluation les composantes de la recherche qui apparaissent périphériques mais sont fondamentales : épistémologie, retombées technologiques, culture scientifique, citoyenneté ? Peut-on imaginer que d'autres acteurs que les chercheurs participent au processus de l'évaluation scientifique ?

7. Faut-il évaluer la science pour la rendre innovante ? À quoi savons-nous que la science progresse ? Y a-t-il d'autres moyens que l'évaluation pour s'assurer que les organisations scientifiques respectent le « contrat éthique de connaissance » passé avec la société et font bon usage des moyens que celle-ci leur confie ? L'intervention d'acteurs extérieurs, gouvernements ou citoyens, conduit-il inévitablement à un dangereux contrôle social sur la recherche ?

8. Comment peut-on intégrer la réflexivité critique sur l'évaluation dans l'évaluation elle-même ? Quelles recommandations peut-on formuler pour que l'évaluation des travaux de recherche et de la vie scientifique maximise l'innovation et la production de connaissances ?

9. …Et, à propos, peut-on faire des suggestions pour améliorer le dispositif d'évaluation d'EspacesTemps.net ?

Vous êtes invité-e à proposer, *jusqu'au 1er mars 2006*, des textes pour publication dans une « Traverse » d'EspacesTemps.*net*. Ces textes peuvent être des récits d'expérience ou des analyses descriptives aussi bien que des énoncés théoriques ou des recommandations. Dans un premier temps, vous pouvez dès maintenant envoyer à Redaction@EspacesTemps.net une proposition, sous la forme d'un résumé de 2000 signes. Vous pouvez également suggérer un mode d'évaluation du texte que vous nous adresserez.

---

Fragen hinsichtlich der wissenschaftlichen Evaluierung.

Evaluierung hat in der wissenschaftlichen Produktion eine zentrale Rolle eingenommen. Standardmäßig bezeichnet der Begriff „Evaluierung" die Beurteilung einer eben abgeschlossenen Arbeit, die mit Hilfe von außerhalb der Erkenntnisziele liegenden Logiken und eines als unabhängig anerkannten Apparats durchgeführt wird. In der Physik, der Biologie, den technologischen Disziplinen und Ingenieurwissenschaften, in einem geringeren Ausmaße, auch in der Mathematik sowie in den Bereichen der Sozialwissenschaften, die versuchen, sich vom Modell der „harten Wissenschaften" zu inspirieren, hat sich in den letzten Jahrzehnten ein „Standardsystem" entwickelt, das sich als ebenso einfach in seinem Prinzip wie folgereich in seinen Konsequenzen zeigt. Dieses System, sehr anschaulich am Beispiel der Zeitschriftenartikel, stellt sich als eine durchaus durchschaubare Prozedur dar: jeder Artikel wird von Referees beurteilt, die von den jeweiligen Verantwortlichen unter den Spezialisten des jeweiligen wissenschaftlichen Spezialgebiets ausgesucht werden. Die Entscheidung, einen Artikel zu publizieren oder nicht, hängt somit in einer mechanischen Weise vom Urteil des Referees ab.

Man kann die Hypothese formulieren, dass der allmählich erreichte Erfolg dieser Methode eine Hierarchisierung der Zeitschriften nach ihrer Legitimität zur Folge gehabt hat. Dies hätte dann einen Effekt von positiven Rückwirkungen ausgelöst, die sämtliche wissenschaftliche Produktionen in ein kohärentes System einbinden. Man kann diese Logik in der folgenden Weise lesen: Zunächst definiert im Rahmen der Publikationen ein als kompetent und gerecht anerkanntes Textbewertungssystem *ipso facto* die Qualität einer Zeitschrift, die es in seinem Betrieb einbezieht. Die Zeitschriften selbst werden auf dieser Basis

in gleicher Weise durch andere Bewerter beurteilt. Das Vorkommen von bibliographischen Referenzen eines Artikels in anderen Artikeln, gemessen durch den *Science citation index*, liefert ein quantifizierbares Mittel, Texte auf Grund ihres Bekanntheitsgrades und ihrer Marktdurchdringungsfähigkeit (*impact factor*) zu ordnen.

Die Evaluierung der Forscher, der Forscherteams und, darüber hinaus, der Institute, der Fakultäten und der Universitäten kann dann in unmittelbarer Weise vom Ausmaß ihres Publizierungs- oder Zitierungserfolgs in den bestrangigen Zeitschriften abgeleitet werden, was automatisch eine private oder öffentliche Finanzierung der betroffenen Akteure zur Folge haben kann, welche ihrerseits ihre Kapazität steigert, weitere Resultate zu produzieren, die wiederum durch den Evaluierungsapparat prämiert werden.

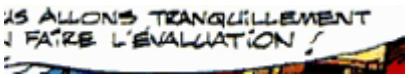
Wegen seines Erfolgs hat dieses System verschiedene negative Einflüsse auf sämtliche Elemente, die ihm nicht angehören: abgewiesene Artikel, Zeitschriften die diese Evaluierungsnormen nicht anwenden, Bücher, unklassifizierbare Veröffentlichungen und, darüber hinaus, Forscher- und Forscherteamtätigkeiten, die nicht nach diesen Normen publizierbare Produkte liefern… Die Evaluierung, aufbauend auf der Publikation von Zeitschriftenartikeln, erscheint so als die Organisationskraft der wissenschaftlichen Legitimität der ganzen Forschungswelt.

Diese Situation ist für Forscher der Sozialwissenschaften in doppelter Hinsicht faszinierend. Erstens, weil sie selbst immer mehr davon betroffen sind; zweitens, weil sie ein an sich interessantes Studienobjekt für ihre eigenen Forschungen darstellt.

In diesem Kontext öffnet EspacesTemps.*net* die Möglichkeit, Beiträge zum Thema einzureichen:

## Eine wissenschaftliche Evaluierung’

### Kann die Dynamik des Wissens normiert werden’




Diese allgemeine Fragestellung wird in einer Serie von Fragen behandelt:
1. Wer bewertet was und wie? Über das Urteil von eingereichten Texten hinaus kann man die Bewertung der Wissenschaftlichen Dimension sämtlicher Produkte in Frage stellen: Doktorarbeiten, wissenschaftliche Bücher, für ein breiteres Publikum bestimmte Abhandlungen, Forschungsprojekte, Forschungsteams und -anstalten, Produkte die nur indirekt der wissenschaftlichen Welt angehören (Philosophie, Universitäts- und Schulhandbücher, Architektur- und Urbanistikprojekte), Tätigkeit der Forscher außerhalb der Forschung… In diesem Sinne kann die wissenschaftliche Komponente jeder sozialen Produktion in Frage gestellt werden.

2. Was sind die Stärken und Schwächen des „Standardsystems“? Erlaubt es einen bedeutsamen Gewinn an Unabhängigkeit der Forscher und ihrer Innovationskraft? Entsteht wegen der Last der Norm und der institutionalisierten Hierarchien ein Risiko, dass die originellsten Beiträge ignoriert werden und die Interdisziplinarität behindert wird? Gibt es Funktionsstörungen in diesem System, die es von seinem eigentlichen Ziel ablenken? Kann der heutige Zeitschriftenmarkt den Pluralismus garantieren oder tendiert er im Gegenteil, seiner oligopolistischen Struktur wegen, zu einer ungerechtfertigten Durchsetzung der *Doxa*? Kann es eine Vereinbarkeit geben zwischen dem standardmäßigen System und einer dynamischen Verlagspolitik der Zeitschriften’

3. Kann man, unabhängig von Kriterien der Wissenschaftlichkeit und deren Veränderungen, Evaluierungsnormen festlegen? Oder soll man im Gegenteil zuerst erstere bestimmen, um daraus Evaluierungsmodi abzuleiten? Kann man eher Fragen bewerten als Antworten, Hypothesen eher als Ergebnisse, globale Theorien eher als einzelne Fortschritte’

4. Gibt es eine Spezifizität der Human- und Sozialwissenschaften bezüglich Evaluierung und Evaluierbarkeit? Wenn ja, woraus besteht sie und was bestimmt sie? Nähern wir uns einer Konvergenz der Evaluierungsmodalitäten in den verschiedenen Forschungsbereichen und –fächern’

5. Kann ein wissenschaftliches Evaluierungssystem mit der Hilfe von Verfahren, die sich auf seinen Apparat beziehen, oder mit der Hilfe von Analyse, welche auf dessen Auswirkungen auf wissenschaftliche Resultate zielen, bewertet werden'

6. Gibt es Alternativen zum Standardsystem? Gibt es andere Modi der Zeitschriftenproduktion (siehe, zum Beispiel, das *Public Library of Science* Projekt), andere Arten von Evaluierungsinstanzen außerhalb der akademischen Welt, die unsere Aufmerksamkeit verdienen? Ändert das Aufkommen anderer Medien der wissenschaftlichen Publikation — wie zum Beispiel des Internets — die Beziehung zwischen Evaluierern und Evaluierten? Kann man während der Evaluierung Forschungskomponenten in Betracht ziehen welche zwar peripher scheinen, jedoch fundamental sind: Epistemologie, technologische Konsequenzen, Wissenschaftskultur, Staatsbürgerschaft? Ist es vorstellbar, dass andere Akteure als Forscher zum Evaluierungsprozess beitragen?

7. Muss die Wissenschaft überhaupt bewertet werden, um innovativ zu sein? Was erlaubt uns zu erkennen, dass die Wissenschaft fortschreitet? Gibt es andere Mittel als die Evaluierung um sicherzustellen, dass die Wissenschaft den mit der Gesellschaft abgeschlossenen „ethischen Wissensvertrag" respektiert und dass von den Mitteln, die sie von ihr zur Verfügung gestellt erhält, guten Gebrauch macht? Leitet das Eingreifen Außenstehender — Akteure, Behörden, Bürger — unvermeidbar zu einer gefährlichen Überwachung der Forschung'

8. Wie kann man die kritische Reflexivität über die Evaluierung in die Evaluierung einbauen? Welche Empfehlungen können formuliert werden, damit die Evaluierung von Forschungsarbeiten und des wissenschaftlichen Lebens mithilft, die Innovation und die Produktion von Wissen zu maximieren'

9. …Und übrigens, gibt es Vorschläge um den Evaluierungsapparat von EspacesTemps.net zu verbessern'

Sie sind dazu eingeladen, *bis zum 1. März 2006* ihre Textvorschläge für eine Veröffentlichung in einer „Traverse" von EspacesTemps.*net* einzuschicken. Ihr Text kann sowohl ein Erfahrungsbericht wie eine beschreibende Analyse sein, eine Theorie wie eine Empfehlung. Sie können ab jetzt einen Textvorschlag an Redaction@EspacesTemps.net in der Form einer 2000 Zeichen langen

Zusammenfassung schicken. Sie können uns ebenfalls einen Evaluierungsmodus des uns zugeschickten Texts vorschlagen.

Picture, Image, Bild : © Éditions Albert René.

Deadline for submission : March 1st 2006.

## Résumé

Questioning scientific evaluation.Evaluation has taken on a major role in scientific production. By default, what one calls ‘evaluation’ mostly refers to the quality assessment of just-produced works — an assessment carried out within an evaluative structure considered as independent of logics exterior to the interests of knowledge. In physics, biology, in technical disciplines and, to ...