Les transformations de l'évaluation de la recherche

Yves Gingras
Chaire de recherche en histoire et sociologie des sciences

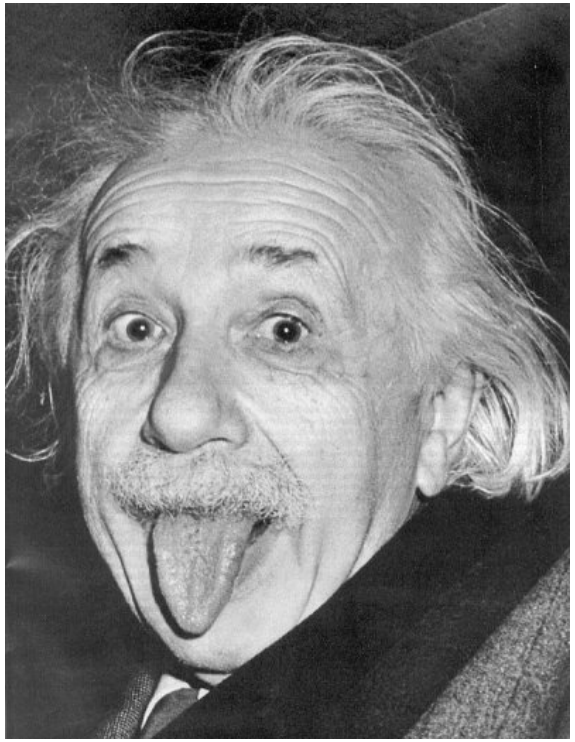
UQAM

Évaluation des publications

Newton évalué par Robert Hooke en 1672:

Hooke critique sévèrement le papier de Newton sur la composition de la lumière soumis aux *Transactions of the Royal Society*; Newton, furieux refuse désormais de publier dans la revue…

Einstein découvre l’évaluation anonyme...



- 1936; papier soumis à *Physical Review*
- Critiqué par un évaluateur anoynyme! Retrait de l’article!
- MAIS...
- Il tient compte des remarques et le publie ailleurs....

La sur-évaluation des enseignants-chercheurs

Evaluation des enseignements

Evaluation des publications

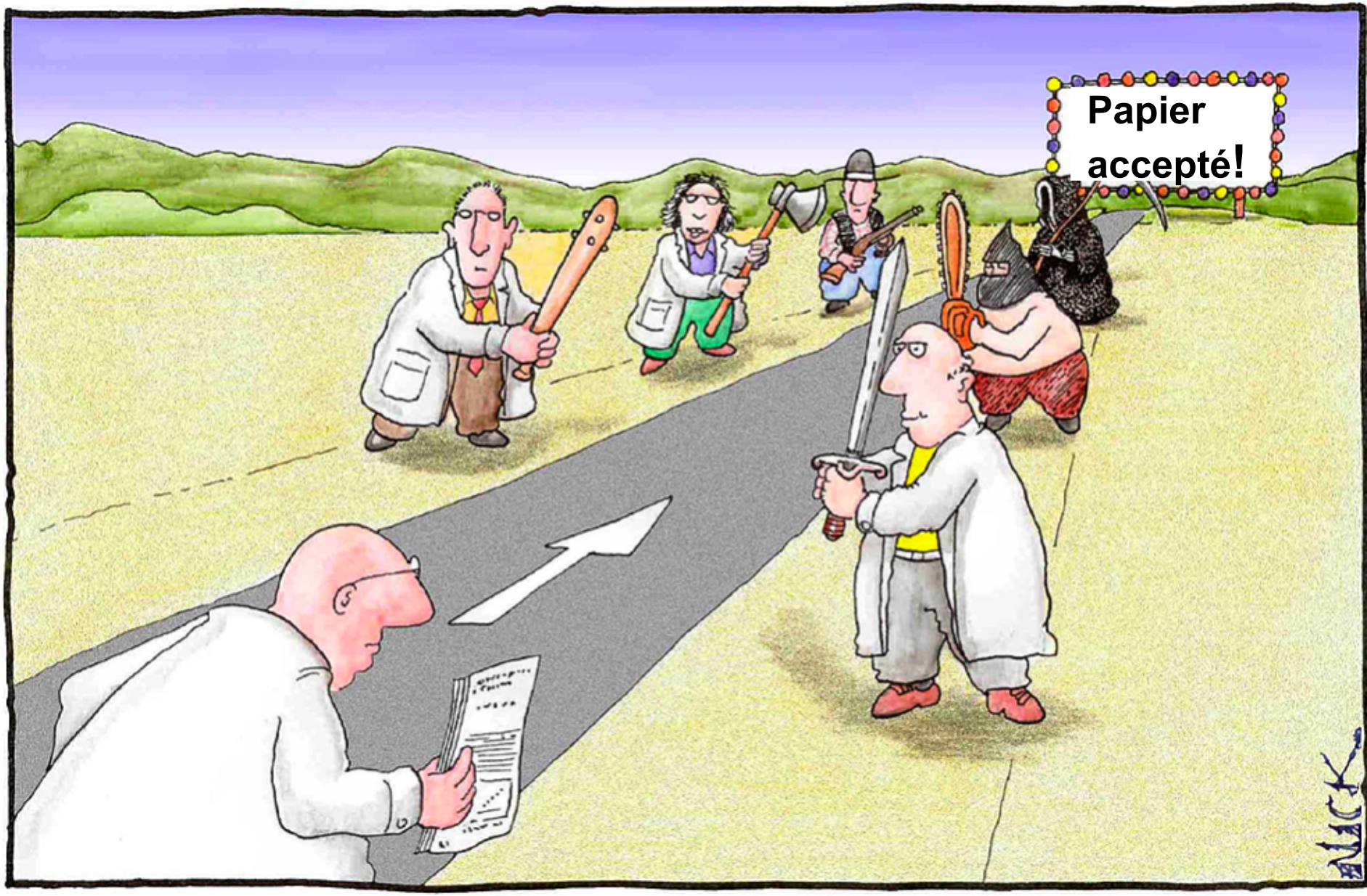
Evaluation des promotions

Evaluation des demandes de subventions

Evaluation des départements et programmes

Evaluation des revues savantes

Evaluation des universités


Papier accepté!

Transformation des mécanismes d'évaluation

- Dominance des comités de pairs jusqu'aux années 1980
- Montée récente des mesures quantitatives: bibliométrie

Pourquoi ?

Citation Analysis: A New Tool for Science Administrators

A new way to assess scientific productivity is about to come into its own. Citation analysis, hitherto an arcane tool of historians and sociologists of science, has now been refined to the point where it offers increasingly interesting possibilities to the science administrator. Proponents of the technique believe that within a few years it will find major uses in decisions at the level of national science policy, as an adjunct of the peer review process, and in evaluating the performance of individual scientists.

ments and as a safety net to catch chemists who write bad grant proposals but are heavily cited.

- Citation analysis has been used in a National Academy of Sciences study showing that postdoctoral fellows supported by the National Institute of General Medical Sciences are more often cited than comparable scientists without postdoctoral training.
- The suspicion that government science policy toward universities is dominated by a coterie of "Eastern Establishment" scientists has been tested

buddy system); moreover it seems likely at present showing to corroborate the decisions made by the embattled peer review system. The NSF has indicated receptivity to a scheme for setting up a citation analysis data bank at a cost of some $250,000 a year.

Articles, of course, are often cited for reasons other than their importance, influence, originality, or merit, and every author can think of a dozen such reasons why citation analysis should not work. He may cite a paper because it was written by his professor, or because he disagrees with it, or because it describes some useful method. But to citation analysts, such references are simply the "noise" in the system, which they believe they can for many purposes filter out or at least reduce to insignificance. In most, but not all, attempts at validation, citation counts are found to correlate highly with almost every conventional measure of scientific quality.

Science, 2 mai 1975

Pourquoi ?

The Controversy over Peer Review

Recent studies of the peer review system show that its critics have yet to make their case.

Thane Gustafson

One of the most striking features of federal support of fundamental science is heavy reliance on scientists outside the government and on the decentralized decision-making processes of the scientific community. Much of the overall direction of federal support for fundamental science depends on the individual decisions of working scientists about which fields to enter and which lines of investigation to pursue. These decisions are communicated to

The most influential advisory committees and the most highly developed system of peer review are to be found in the NIH, which funds 75 percent of the biomedical research performed in American medical schools as well as more than 40 percent of all university research. The NIH peer review mechanism has two parts. Initial review of proposals for research grants is delegated to approximately 50 study sections composed entirely of nongovernment

mail review. Single-stage review is used by certain divisions of the NSF (3). The initial evaluation of proposals is done by NSF program directors, who make a preliminary ranking. Although advisory panels of outside experts exist for each NSF program, their role is confined mainly to a review of cases considered borderline by the program director. The outside experts do not assign priority scores, and the final assignment of priorities is left to the program director. Outside scientists are not without influence, however; in some NSF programs the influence of the panels is greater than that of the program director (4).

Mail review, which is also used in NSF as well as in other federal agencies, is the form of peer review that gives the greatest role to the program director. Project proposals are sent by mail to outside reviewers around the country, but the opinions of the reviews are not binding on the program director. Even in this form of peer review, however, the influence of the program directors is limited by the fact that all large projects are subject to the approval of the National Science Board (5).

Science, 12 décembre 1975

Pourquoi?

- Recherche de mesures « objectives »
- Contrôle de l'expertise des « pairs »
  - Passage des « pairs » aux « experts »
- Opérations sur des nombres: commensurabilité, moyennes, classements en rangs, etc

CRITERION 1: QUALITY OF THE PRINCIPAL INVESTIGATOR

Score: 1.0 (Non-competitive) 1.5 2.0 (Very Good) 2.5 3.0 (Excellent) 3.5 4.0 (Outstanding) Above Threshold

Intellectual capacity and creativity:
To what extent are the achievements and publications of the Principal Investigator ground-breaking and demonstrative of independent creative thinking and capacity to go significantly beyond the state of the art?

To what extent will an ERC Starting Grant make a significant contribution to the establishment or consolidation of independence?

Commitment:
Is the Principal Investigator strongly committed to the project and willing to devote a significant amount of time to it (they will be expected to devote at least 50% of their working time to the ERC-funded project)?

Comments: (minimum 50 - maximum 4000 characters)
...

La fièvre de l’évaluation

- Multiplication des classements
  - US News and World Report
  - Classement de Shanghai
- Multiplication des indicateurs
  - Facteur d’impact des revues
  - Indice h
  - Indice K…

OP-ED COLUMNIST

The College Rankings Racket

By JOE NOCERA

Published: September 28, 2012 | 551 Comments

The U.S. News & World Report's annual college rankings came out earlier this month and — knock me over with a feather! — Harvard and Princeton were tied for first.

Nouveau classement!

Ranking By Campus Setting : Category

Type of Higher Education Institution : Category

World Ranking 2011

University	Country	Ranking	Total Score	Setting And Infrastructure	Energy and Climate Change	Waste	Water	Transportation
University of Nottingham		1	**8,033.54**	1,394.54	2,534.00	1,275.00	1,155.00	1,675.00
Northeastern University		2	**7,981.46**	1,396.46	2,370.00	1,350.00	1,440.00	1,425.00
University of Connecticut		3	**7,708.02**	1,458.02	2,100.00	1,125.00	1,450.00	1,575.00
University College Cork		4	**7,682.48**	1,336.48	2,126.00	1,275.00	1,270.00	1,675.00
Linkoping University		5	**7,661.23**	1,406.25	2,104.98	1,275.00	1,200.00	1,675.00
University of California, Berkeley		6	**7,601.87**	1,279.37	2,220.00	1,425.00	1,127.50	1,550.00

Outil de marketing…

La Tribune, 29 janvier 2014, page 3

DÉVELOPPEMENT DURABLE

L'UdeS conserve la première position au Québec et au Canada

ISABELLE PION
isabelle.pion@latribune.qc.ca

SHERBROOKE — Pour une deuxième année consécutive, l'Université de Sherbrooke obtient la première position en matière de développement durable au Québec et au Canada, selon le UI GreenMetric World University Ranking (WUR), un palmarès produit par l'Université d'Indonésie. Le palmarès la place 6e au monde parmi les universités qui participent à ce classement, soit 301 établissements universitaires à travers le monde.

L'UdeS avait obtenu le même classement l'an passé. Elle a toutefois obtenu 151 points de plus que l'année dernière, pour un total de 6948 points. À l'échelle canadienne, une dizaine d'universités se retrouvent dans ce classement, dont les universités Concordia et York.

Le GreenMetric WUR est publié pour la quatrième fois. Il compare les mesures prises par les universités pour favoriser le développement durable de leurs campus et s'intéresse à leur gestion des enjeux environnementaux. Il se penche sur plusieurs critères, comme la gestion des matières résiduelles, la gestion de l'eau, le transport durable, etc.

C'est l'Université de Nottingham, au Royaume-Uni, qui a obtenu la première place, avec un total de 7521 points.

L'UdeS obtient la meilleure note accordée (1000 points) pour sa gestion de l'eau. L'institution a mis en place plusieurs mesures au cours des dernières années, dont l'installation d'appareils à faible débit et l'interdiction de l'arrosage extérieur. Plusieurs initiatives ont aussi été lancées afin d'améliorer la gestion des matières résiduelles, comme

accès au transport en commun offert aux étudiants, l'introduction de véhicules électriques et de bornes de recharge sur les campus, de même que des mesures favorisant le transport actif.

« Nous constatons avec grand plaisir que les mesures mises en place à l'UdeS permettent de se maintenir parmi les universités les plus innovantes à l'échelle internationale », a commenté par voie de communiqué Alain Webster, vice-recteur au développement durable, aux relations gouvernementales et au campus de Longueuil. La stratégie de développement durable, maintenant bien ancrée dans notre communauté universitaire, constitue un élément important de notre plan stratégique et contribue à faire de nos campus de véritables laboratoires de développement durable au bénéfice de l'ensemble de la communauté. »

Évaluer les évaluations...

- Le nouveau management public considère trois grand principes:
- Reddition de comptes
- Efficience
- Transparence
- Il faudrait les appliquer à... l'évaluation!

REF 2014 cost almost £250 million

Accountability review finds cost of assessment equates to 2.4 per cent of funding bodies' expected spend over next six years

July 13, 2015

Tout cela pour conclure que :
-75% des universités Britanniques sont "excellentes au plan international (46%)" ou "leader mondial" (30%) et 20% étant même "reconnues au plan international"!!

Exemples d’indicators non valides

- L’indice h et ses dérivés
- Facteur d’impact de revues appliqués aux articles
- Classements des universités
- Mesures fondées sur internet: twitter…

Effets pervers

Transforme l'article **d'unité de connaissance** en « **unité comptable** »

Usage stratégique des citations

Choix des objets de recherche les plus « vendables » aux revues bien classées…

Multiplication des articles

Émergence de revues « prédatrices »

Détourne les institutions des priorités locales

Achat des auteurs et de leurs adresses

Conclusion

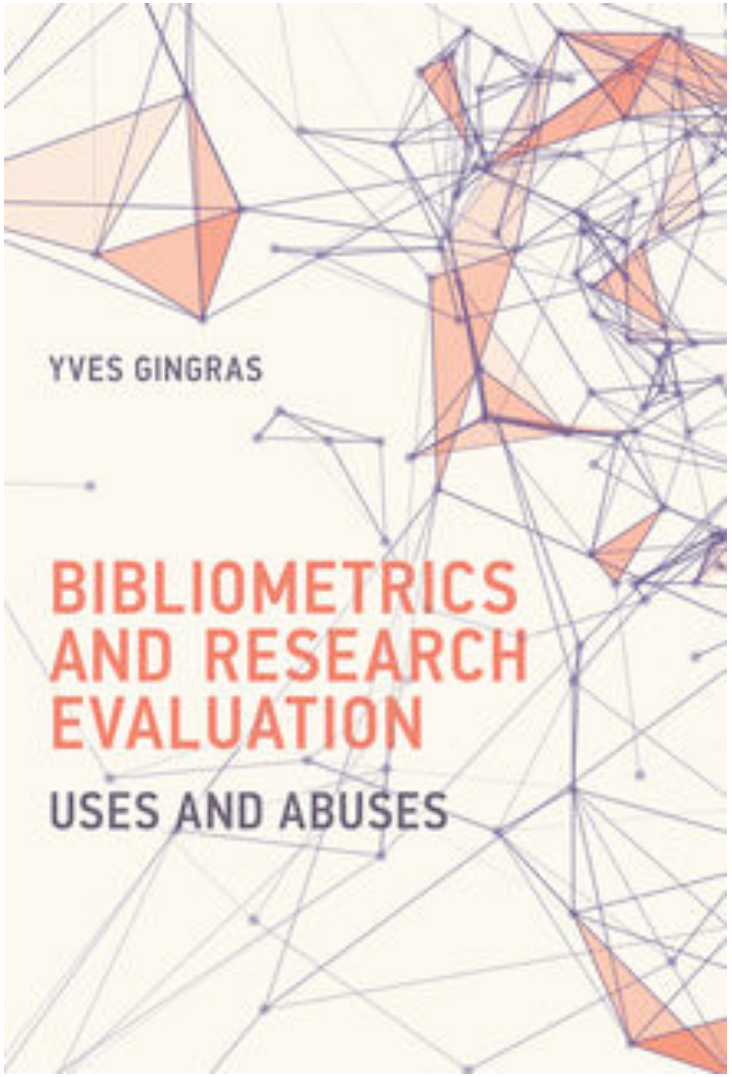
- Adapter les techniques d'évaluation à l'échelle d'analyse
- S'assurer que les indicateurs de performance choisis soient valides, multiples et correspondent à la mission de l'institution évaluée.
- Les chercheurs et les gestionnaires de la recherche doivent éviter de donner de la crédibilité à des indicateurs **improvisés et inadéquats à leur objet**.
- La mission des institutions de recherche doit tenir compte des besoins spécifiques du pays et les indicateurs doivent être choisis en fonction de la mission spécifique et **non pas importés et copiés en fonction de modes** soi-disant « internationales ».
- Limiter la montée du « marketing » et des « comm » au sein des universités, pour qui toute « pub » est bonne…

Morale…

« Tout ce qui <u>peut être compté</u> ne compte pas nécessairement; tout <u>ce qui compte</u> ne se compte pas nécessairement »

Citation attribuée à Albert Einstein

Pour plus de détails….



ou

…Merci