CREMA

Center for Research in Economics, Management and the Arts

Academic Rankings and Research Governance

Working Paper No. 2010 - 04

CREMA Gellertstrasse 18 CH - 4052 Basel www.crema-research.ch

Academic Rankings and Research Governance

(this version 11.3.2010)

by

Margit Osterloh[a,c,*], Bruno S. Frey[b,c]

[a]*University of Zurich*

Institute of Organization and Administrative Sciences

Universitaetsstrasse 84

CH-8006 Zurich, Switzerland

Email: osterloh@iou.uzh.ch

[b]*University of Zurich*

Institute for Empirical Research in Economics

Winterthurerstrasse 30

CH-8006 Zurich, Switzerland

Email: bsfrey@iew.uzh.ch

[c]*CREMA - Center for Research in Management, Economics and the Arts, Zurich*

Gellertstrasse 18, CH-4052 Basel, Switzerland

Corresponding address: Prof. Dr. Dr. h.c. Margit Osterloh, Institute of Organization and Administrative Science, Universitaetsstrasse 84, CH-8006 Zurich, Switzerland.
Tel.: +41 44 634 28 41; fax: +41 44 634 49 42
E-mail: osterloh@iou.uzh.ch.

Abstract

Academic rankings today are the backbone of research governance, which seem to fit the aims of "new public management" on the one side and the idea of the "republic of science" on the other side. Nevertheless rankings recently came under scrutiny. We discuss advantages and disadvantages of academic rankings, in particular their unintended negative consequences on the research process. To counterbalance these negative consequences we suggest (a) rigorous selection and socialization, and (b) downplaying the impact of rankings in order to reconcile academic self-governance with accountability to the public.
(89 words)

JEL Classification No: H52, H83, I23, J44, L38

Key words: peer reviews, rankings, research governance, psychological economics, new public management, economics of science, control theory.

Introduction

Academic rankings today are generally considered the backbone of research governance in academia. On the one hand they are based on the evaluation of scientific peers who are the only ones able to estimate the quality of research. On the other hand rankings are considered to give the public a transparent picture of scholarly activity and to make universities more accountable for their use of public money. They are intended to unlock the "secrets of the world of research" (Weingart 2005, p. 119) for journalists as well as for deans, administrators, and politicians who have no special knowledge of the field. They provide a basis for control, for the allocation of resources and for the provision of compensation packages (e.g. Worrell 2009). However, in recent times peer reviews and academic rankings have come under scrutiny. A lively discussion about the quality of peer reviews (e.g. Lawrence 2002, 2003; Frey 2003; Starbuck 2005, 2006; Abramo, Angelo, and Caprasecca 2009) and academic rankings (e.g. Adler and Harzing 2009; Albers 2009) takes place. This discussion focuses mainly on issues of method and how to improve it. It is taken for granted that more and better indicators are needed to enhance the quality of rankings (e.g. Starbuck 2009). Only in a few cases it is asked whether the advantage of controlling research activities from outside may

produce unintended negative side effects, even if indicators for research quality were perfect (e.g. Weingart 2005; Espeland and Sauder 2007). As a consequence, the question has not been raised whether there are viable alternatives to academic rankings as an instrument for academic governance.

In this article, we discuss two issues. First, we analyze the advantages and disadvantages of rankings, in particular their unintended negative consequences. Second we ask whether there exists an alternative to academic rankings as the main instrument of academic governance.

We begin by analyzing two conceptual pillars of rankings as the basis of our present research governance, namely on the one side "new public management" and on the other side the concept of the "republic of science". The second section presents empirically based findings on the advantages and disadvantages of rankings and suggestions made on how to overcome their shortcomings. The third section focuses on an aspect mostly disregarded, namely the behavioral reactions to rankings which may overcompensate their advantages. The last section discusses whether and to which extent there are viable alternatives to rankings as the dominant instruments of research governance.

Conceptual Issues: "New Public Management" versus "Republic of Science"?

Over the past years, universities have increasingly adopted the idea of "new public management", namely the idea that universities, like other public services such as hospitals, schools or public transport should be subjected to a similar governance as for-profit enterprises. "More market" and "strong leadership" have become the keywords (Schimank 2005). This is reflected in procedures transferred from private companies like management by objectives or pay-for-performance for scholars. Overall, the reforms are aimed at the establishment of an "enterprise university" (e.g., Clark 1998; Marginson and Considine 2000; Bok 2003; Khurana 2007; Donoghue 2008). A number of processes have been identified as drivers behind this development (e.g. Bleiklie and Kogan 2007; Schimank 2005).

First, the rise of mass education during the 1980s and 1990s made higher education more expensive and visible to the public. This fact contributed to pressure for efficiency and accountability towards the tax-payer.[1] Second, it has been criticized that the traditional system of self-governance in universities has impeded the necessary reforms towards mass education. New public management was seen as a way of breaking the "reform blockade". Third, a growing demand for relevance of research became influential in the public debate. In their

[1] For an overview of the transition to mass higher education in various countries see Teichler (1988).

book “The New Production of Science” Gibbons et al. (1994) claimed that science has been transformed from a traditional university- and discipline-centered “Mode 1” knowledge production to a so called transdisciplinary “Mode 2” knowledge production in which stakeholders from outside the university are involved.[2] Therefore criteria of quality are no longer determined by academic peers only. Research comes under pressure to legitimate its outcomes to people outside academia. Fourth, “economics has won the battle for theoretical hegemony in academia and society as a whole“ (Ferraro, Pfeffer, and Sutton 2005, p. 10). As a consequence, standard economics, in particular the principal agent view, has gained dominance not only in corporate governance (Daily et al. 2003) but also in public and academic governance. According to standard economics, scholars have to be monitored and sanctioned in the same way as other employees. The underlying assumption is that control and correctly administered pay-for-performance schemes positively impact motivation and lead to an efficient allocation of resources (Propper 2006). Taken together, the ideals about the governance of universities have changed from a “republic of scholars” to a “stakeholder organization” in which the voice of scholars is but one among several stakeholders and professorial autonomy is curtailed (Speckbacher et al. 2008).

This view stands at first glance in stark contrast to the ideal of self-governance of the scientific community.[3] This ideal was undisputed for a long time. Over three hundred years ago, Gottfried Leibniz promoted the “republic of letters” – an independent, self-defining network of scholars that transcends national and religious boundaries (Leibniz 1931).[4] Polanyi (1962/2002, p. 479) contends “The soil of academic science must be exterritorial in order to secure its rule by scientific opinion.” His “republic of science” is based on the self-co-ordination of independent scientists. Authority “is established between scientists, not above them.” (p. 471). Authors like Bush (1945), Merton (1973), and Stokes (1997) warn that outside actors are tempted to shape science according to their own value systems and thus jeopardize the mission of science. This view is supported by the economics of science (Arrow 1962; Nelson 1959, 2004; Dasgupta and David 1994; Stephan 1996). According to this view, in academia the evaluation by peers has to substitute for the evaluation by the market because of two fundamental characteristics of science, its public nature and its high uncertainty. The public nature of scientific discoveries which leads to a market failure has been intensively discussed by Arrow (1962) and Nelson (1959, 2006). The fundamental uncertainty of scientific endeavors also leads to a market failure. It exists because success in academia is

[2] For a criticism of this approach see Weingart (1997).

[3] As Lawrence (2003, p. 259) puts it “Managers are stealing power from scientists”.

[4] For a discussion see Ultee (1987).

reflected by success in the market often only after a long delay or sometimes not at all (Bush 1945; Nelson 1959, 2004, 2006). In addition, research often produces serendipity effects; that is, it provides answers to unasked questions (Stephan 1996; Simonton 2004). As it is often not predictable which usefulness a particular research endeavor produces and whether it ever will be marketable, peers instead of the market have to evaluate whether a piece of research represents an advance. Peers have the opportunity to identify possible errors and risks; they can profit themselves from the innovation to push forward their own research; redundancies are avoided; and the new knowledge can quickly be used for new and cheaper technologies. Due to failure of markets and prices there is a special "currency" that governs the republic of science, the priority rule (Merton 1957; Dasgupta and David 1994; Stephan 1996; Gittelman and Kogut 2003). This rule attributes success to the person who first makes an invention, and who the scientific community recognizes to be first. The priority rule serves two purposes, hastening discoveries, and hastening their disclosure (Dasgupta and David 1994, p. 499): A discovery must be communicated as quickly as possible to the community of peers in order to gain their recognition.

Consequently, the peer review system is taken to be the founding stone of academic research evaluation. Indicators are awards, honorary doctorates, or membership in prestigious academies (Stephan 1996; Frey and Neckermann 2008).[5] Its main form for the majority of scholars consists of publications and citations in professional journals with high impact factors. Such indicators are provided by academic rankings, based on peer-reviewed publications, citations, and the impact factors of journals like Thomson Reuters's Impact Factor (JIF) (see Garfield 2006, for a historical review) and the relatively recent h-index (Hirsch 2005).[6]

In that view, a well-designed governance system based on academic rankings seems to combine perfectly an output-oriented evaluation of researchers, as postulated by new public management, with the requirements of a peer-based evaluation system, as postulated by the economics of science. It is based on the one side on evaluations of the peers who are able to assess the quality of research from inside the scientific world. On the other side it seems to be an easy to understand measure for non-experts like politicians, administrators and other stakeholders to evaluate the quality of research from outside. Therefore, today these measures are adopted almost universally in academia for most things that matter as part of the present

[5] Zuckerman (1992) estimates that by the beginning of the 1990s around 3,000 different scientific awards existed in North America.

[6] Examples of prominent rankings are ISI Web of Knowledge Journal Citation Report ; ISI Web of Knowledge Essential Science Indicators ; IDEAS Ranking ; Academic Ranking of World Universities ; or Handelsblatt Ranking .

research governance system: tenure, salary, grants, and budget decisions. This has lead to an ever growing evaluation industry and actively marketed tools like the ISI Web of Science.

Empirically Based Findings on Academic Rankings

Academic rankings have become prominent because of two reasons. First, they are intended to give the public an overview over the success of research activities. Second, they avoid some problems of qualitative peer reviews which have been discussed recently (e.g., Armstrong, 1997; Wenneras and Wold 1999; Brook 2003; Frey 2003; Bedeian 2004; Starbuck 2005, 2006; Tsang and Frey 2007; Gillies 2005, 2008; Abramo et al. 2009):[7]

- Low inter-rater reliability. There is an extensive literature on the low extent to which reviewing reports conform to each other (Miner and MacDonald 1981; Cole 1992; Weller 2001). The correlation between the judgments of two peers falls between 0.09 and 0.5 (Starbuck 2005). A much discussed study of peer reviewing was conducted by Peters and Ceci (1982). They resubmitted 12 articles to the top-tier journals that had published them only 18 to 32 months earlier, giving the articles fictitious authors at obscure institutions. Only three out of 38 editors and reviewers recognized that the articles had already been published. From the remaining nine articles, eight were rejected. It is important that the correlation is higher for papers rejected than for papers accepted (Cichetti 1991). This means that peer reviewers are better able to identify academic low performers; that is, it is easier to identify papers that do not meet minimum quality standards than those that are a result of excellent research (Lindsey 1991; Moed 2007).
- Low prognostic quality. The reviewers' rating of manuscripts quality is found to correlate only 0.24 with later citations (Gottfredson 1978). According to Starbuck (2006, pp. 83–84), the correlation of a particular reviewer's evaluation with the actual quality as measured by later citations of the manuscript reviewed is between 0.25 and 0.3. This correlation rarely rises above 0.37, although there is evidence that higher prestige journals publish more high-value articles (Judge, Cable, Colbert, and Rynes 2007). Because of some randomness in editorial selections (Starbuck 2005),[8] one editor even advises rejected authors to "Just Try, Try Again" (Durso 1997).[9]

[7] See also the special issue of *Science and Public Policy* (2007) and the Special Theme Section on "The use and misuse of bibliometric indices in evaluating scholarly performance" of *Ethics in Science and Environmental Politics*, 8 June 2008.

[8] See also the "Social Text"-Affair, which deals with the malfunction of editors: The physicist Alain D. Sokal published an article in a (non refereed) special issue of the journal "Social Text" which was written as a parody. The editors did not realize the bogus article as a hoax, see Sokal (1996).

[9] However, this strategy overburdens reviewers and may lower the quality of reviews. For example, they have neither enough time nor the incentive to check the quality of the data and of the statistical methods employed, as

- Low consistency over time. Many rejections of papers in highly ranked journals are documented that later were awarded high prizes, including the Nobel Prize (Gans and Shepherd 1994; Campanario 1996; Horrobin 1996; Lawrence 2003). This means that in the case of radical innovations or paradigm shifts (Kuhn 1962) peer reviews often fail.
- Confirmation biases. Reviewers find methodological shortcomings in 71 percent of papers contradicting the mainstream, compared to only 25 percent of papers supporting the mainstream (Mahoney 1977).

As a reaction to the criticism of qualitative peer reviewing, bibliometric methods, that is, rankings based on the number of publications, citations, and impact factors have become more prominent.[10] Though rankings are based on qualitative peer reviews it is expected that some of the problems discussed are counterbalanced by the following advantages of rankings (e.g., Abramo et al. 2009).

- Rankings are more objective because they are based on more than the three or four evaluations typical for qualitative approaches. Through statistical aggregation individual reviewers´ biases may be balanced out (Weingart 2005).
- The influence of the old boys' network may be avoided. An instrument is provided to dismantle unfounded claims to fame. Rankings can serve as fruitful, exogenous shocks to some schools and make them care more about the reactions of the public (Khurana 2007, p. 337).
- Rankings are cheaper than pure qualitative reviews, at least in terms of time. They admit updates and rapid intertemporal comparisons.

However, in recent times it became clear that bibliometric measures may counterbalance some problems of qualitative peer reviews, but that they have disadvantages of their own (Butler 2007; Donovan 2007; Weingart 2005; Adler et al. 2008; Adler and Harzing 2009). Until now, mainly technical and methodological problems were highlighted (van Raan 2005).

Technical problems consist of errors in the citing-cited matching process, leading to a loss of citations to a specific publication. First, it is estimated that this loss amounts on average to 7 percent of the citations. In specific situations, this percentage may even be as high as 30 percent (Moed 2002). Second, there are many errors made in attributing

some striking examples in economics demonstrate (Hamermesh 2007).

[10] For example the British Government decided to replace its Research Assessment Exercise based mainly on qualitative evaluations with a system based mainly on bibliometrics. Interestingly, the Australian Government, which has used mostly bibliometrics in the past, plans in the future to strengthen qualitative peer review methods (Donovan 2007).

publications and citations to the source, for example, institutes, departments, or universities. In the popular ranking of the Shanghai Jiao Tong University, these errors led to differences of possibly 5 to 10 positions in the European list and about 25 to 50 positions in the world list (Moed 2002). The impact factor of Thomson´s ISI Web of Science, is accused of having many faults (Monastersky 2005; Taylor, Perakakis, and Trachana 2008). It is unlikely that the errors are distributed equally. Kotiaho, Tomkin, and Simmons (1999) find that names from unfamiliar languages lead to a geographical bias against non-English speaking countries. Third, it has been shown that small changes in measurement techniques and classifications can have large effects on the position in rankings (Ursprung and Zimmer 2006; Frey and Rost forthcoming).

Methodological problems of constructing meaningful and consistent indices to measure scientific output have been widely discussed recently (Lawrence 2002, 2003; Frey 2003, 2009; Adler et al. 2008; Adler and Harzing 2009). Therefore, we briefly mention the main problems discussed in the literature.

First, there are selection problems. Often only journal articles are selected for incorporation in the rankings, although books, proceedings, or blogs contribute considerably to scholarly work. Other difficulties include the low representation of small research fields, non-English papers, regional journals, and journals from other disciplines even if they are highly ranked in their respective disciplines. Hence, collaboration across disciplinary boundaries is not furthered.

Second, citations can have a supportive or rejective meaning or merely a herding effect. The probability of being cited is a function of previous citations according to the "Matthew effect" in science (Merton 1968). Simkin and Roychowdhury (2005) estimate that, according to an analysis of misprints turning up repeatedly in citations, about 70–90 percent of scientific citations are copied from the list of references used in other papers; that is, 70–90 percent of the papers cited have not been read. Consequently, incorrect citations are endemic. They are promoted by the increasing use of meta-analyses, which generally do not distinguish between high and low quality analyses (Todd and Ladle 2008). In addition, citations may reflect fleeting references to fashionable "hot topics."

Third, using the impact factor of a journal as a proxy for the quality of a single article leads to substantial misclassification. It has been found that many top articles are published in non-top journals, and many articles in top journals generate very few citations in management research (Starbuck 2005; Singh, Haddad, and Chow 2007), economics (Laband and Tollison 2003; Oswald 2007), and science (Campbell 2008). A study of the "International

Mathematical Union" even concludes that the use of impact factors can be "breathtakingly naïve" (Adler et al. 2008, p.14) because it leads to large error probabilities.

Fourth, there are difficulties comparing citations and impact factors between disciplines and even between subdisciplines (Bornman et al. 2008).

Implications Discussed to Overcome the Problems of Rankings

In recent times some suggestions have been made to deal with the technical and methodological problems of rankings.

First, a temporary moratorium of rankings is suggested "until more valid and reliable ways to assess scholarly contributions can be developed" (Adler and Harzing 2009, p. 72). As is the case for most authors, they believe that the identification of particular shortcomings should serve as a stepping stone to develop a more reliable research evaluation system (see also Abramo et al. 2009; Starbuck 2009). In contrast, policy-makers admit that indicators like rankings and grants are spurious. But as long as scholars present no better data, they will use it since they believe that the present data are better than none (e.g. Schimank 2005).

Second, it has been argued that bibliometric indicators should not be used as ready-to-go indicators lacking the competence to understand what is being measured (van Raan 2005; Weingart 2005). Therefore, standards of good practice for the analysis, interpretation, and presentation of bibliometric data should be developed and adhered to when assessing research performance. This needs a lot of expertise (Bornmann et al. 2008), which constrains considerably the responsible use of rankings as a handy instrument for politicians, administrators, and journalists to assess academic performance by rankings.

Third, it is suggested to use a number of rankings (e.g. Adler and Harzing 2009), since their results differ markedly, in particular with respect to rankings of individuals (Frey and Rost forthcoming). Again, this suggestion constrains rankings as easy to handle instruments for non-experts.

Fourth, a combination of qualitative peer reviews and bibliometrics, so-called informed peer reviews, could be applied. It is argued that they can balance the advantages and disadvantages of these two methods (Weingart 2005; Butler 2007; Moed 2007).

Fifth, a holistic approach of evaluation has been suggested, which combines measures of research quality and impact with peer and user evaluation, taking into account the views of various stakeholders inside and outside academia (Donovan 2007). However this approach bears the danger of compromising on the smallest common denominator and of inhibiting research with unorthodox or uncertain outcomes.

These suggestions may to some extent mitigate the problems of rankings, but they make the use of rankings difficult for non-experts and thus are not able to reconcile the aims of "new public management" with the "republic of science" as intended. Moreover, even if rankings worked perfectly, they cannot overcome the problems of behavioral reactions to rankings (Osterloh and Frey 2009).

Behavioral Reactions to Rankings

Even if over time the methodological and technical problems could be coped with, severe problems remain, caused by unintended side effects of rankings on the side of individuals and institutions. First they consist in the so-called reactive measures (Campbell 1957), caused by the fact that people change their behavior strategically in reaction to being observed or measured, in particular if the measurement is not accepted voluntarily (Espeland and Sauder 2007). Reactivity threatens the validity of measures according to the saying "When a measure becomes a target, it ceases to be a good measure" (Strathern 1996, p 4). Second the unintended consequences consist in the danger of reducing the intrinsically motivated curiosity of researchers. Both problems, which are discussed only by few authors in the research governance literature, have consequences on the level of individual scholars and on the level of institutions.

Level of individual scholars

Reactivity on the level of individual scholars may take on the one hand the form of goal displacement and on the other hand the form of counterstrategies to "beat the system."

Goal displacement (Perrin 1998) means that people maximize indicators that are easy to measure and disregard features that are hard to measure. This problem is also discussed as the multiple-tasking effect (Holmstrom and Milgrom 1991; Ethiraj and Levinthal 2009). There is much evidence of this effect in laboratory experiments (Staw and Boettger 1990; Gilliland and Landis 1992; Schweitzer, Ordonez, and Douma 2004; Ordonez, Schweitzer, Galinsky, and Bazerman 2009)[11]. For example, Fehr and Schmidt show that output-dependent financial incentives lead to the neglect of non-contractible tasks.

In academia examples can be found e.g. in the "slicing strategy" whereby scholars divide their research results to a "least publishable unit" (Weingart 2005, p. 125) by breaking them into as many papers as possible to increase their publication list. Another example of

[11] Locke and Latham (2009) in a rejoinder provide counterevidence to Ordonez et al. (2009). They argue that goal setting has no negative effects. However, they disregard that goal setting may well work for simple but not for complex tasks within an organization. For the latter case, see Earley, Connolly, and Ekegren (1989) and Ethiraj and Levinthal (2009).

goal displacement is the lowering of standards for PhD candidates when the amount of completed PhDs is used as a measure in rankings. Empirical field evidence of goal displacement in academia is shown in an Australian study (Butler 2003). The mid-1990s saw a linking of the number of peer-reviewed publications to the funding of universities and individual scholars. The number of publications increased dramatically, but the quality as measured by relative citation rates decreased.[12]

Counterstrategies are more difficult to observe than goal displacement. They consist of altering research behavior itself in order to "beat the system" (Moed 2007). Numerous examples can be found in educational evaluation (e.g., Haney 2002; Nichols, Glass, and Berliner 2006; Heilig and Darling-Hammond 2008). The following behaviors are of special relevance in academia.

Scholars distort their results to please, or at least not to oppose, prospective referees. Bedeian (2003) finds evidence that no less than 25 percent of authors revised their manuscripts according to the suggestions of the referee although they knew that the change was incorrect. Frey (2003) calls this behavior "academic prostitution"

Authors cite possible reviewers because the latter are prone to judge papers more favorably that approvingly cite their work, and these same reviewers tend to reject papers that threaten their previous work (Lawrence 2003, p. 260).[13] Authors willingly adapt to editors who pressure them to cite their respective journals in order to raise their impact rankings (Garfield 1997; Smith 1997; Monastersky 2005)

To meet the expectations of their peers—many of whom consist of mainstream scholars—authors may be discouraged from conducting and submitting creative and unorthodox research (Horrobin 1996; Prichard and Willmott 1997; Armstrong 1997; Gillies 2008).

The effects of reactivity are enforced if the second kind of unintended consequences takes place, the decrease of intrinsically motivated curiosity which generally is acknowledged to be of decisive importance in academic research (Amabile 1996, 1998; Stephan 1996; Simonton 2004). There exists considerable empirical evidence in psychology and psychological economics[14] that there is a crowding-out effect of intrinsic motivation by

[12] It could be argued that a remedy to this problem consists of resorting to citation counts. While this remedy overcomes some of the shortcomings of publication counts, it is subject to the technical and methodological problems mentioned.

[13] Such problems of sabotage in tournaments have been extensively discussed in personnel economics, see Lazear and Shaw (2007).

[14] We prefer the expression "psychological economics" rather than the more common expression "behavioral economics" for two reasons. First, economists had already examined human behavior before this new field emerged. Second, Simon points out that the term „behavioral" is misleading since it may be confounded with the „behaviorist" approach in psychology.

externally imposed goals linked to incentives which do not give a supportive feedback and are perceived to be controlling[15] (Frey 1992, 1997; Deci, Koestner, and Ryan 1999; Gagné and Deci 2005; Falk and Kosfeld 2006; Ordonez et al. 2009).[16]

From that point of view rankings tend to crowd out intrinsically motivated curiosity. First, in contrast to qualitative peer reviews rankings do not give a supportive feedback since they do not tell scholars how to improve their research. Second, since rankings are mostly imposed from outside the content of research is in danger of losing importance. It is substituted by the position in the rankings (Kruglansky 1975). As a consequence, the dysfunctional reactions of scholars like goal displacement and counterstrategies are enforced because they are not constrained by intrinsic preferences. The inducement to "game the system" in an instrumental way may get the upper hand.

Level of institutions

Reactivity on the institutional level takes several forms. First, if rankings are used as measure to allocate resources and positions they create a *lock-in effect*. Even those scholars and academic institutions that are aware of the deficiencies of rankings do well not to oppose them. If they did so, they would not only be accused of being afraid of competition, but also of not contributing to the prestige and resources of their department or university. Therefore, it is a better strategy to follow the rules and to play the game. For example, in several countries, highly cited scientists are hired immediately before the evaluation of departments and programs are scheduled to take place in order to raise publication and citation records. Such stars are highly paid although they often have little involvement with the respective university (Brook 2003; Stephan 2008)

Second, a negative *walling-off effect* sets in. Scholars themselves are inclined to apply rankings to evaluate candidates in order to gain more resources for their research group or department. In addition, it is easier to count the publications and citations of colleagues than to evaluate the content of their scholarly contributions. By doing this, scholars delegate their own judgment to the counting exercise behind rankings, although, by using such metrics, they admit their incompetence in that subject (Browman and Stergiou 2008). This practice is defended by arguing that specialization in science has increased so much that even within disciplines it is impossible to evaluate the research in neighboring fields (Swanson 2004; van

[15] A third precondition is social relatedness, see Gagne and Deci (2005).

[16] The crowding-out effect sometimes is contested e.g. Eisenberger and Cameron (1996); Gerhart and Rynes (2003); Locke and Latham (2009). However the empirical evidence for complex tasks and actors intrinsically motivated in the first place is strong, see Deci, Koestner, and Ryan (1999); Weibel, Rost, and Osterloh (2009), for a survey of the empirical evidence, see Frey and Jegen (2001).

Fleet, McWilliams, and Siegel 2000). However, this practice in turn reinforces specialization and furthers a walling-off effect between disciplines and subdisciplines. By using output indicators instead of communicating on the contents, the knowledge in the various fields becomes increasingly disconnected. This hampers the ability to create radical innovations that often cross disciplinary borders (Amabile et al. 1996; Dogan 1999).

Third, *research is increasingly homogenized*. Research endeavors tend to lose the diversity that is necessary for a creative research environment. This consequence was pointed out for business schools by Gioia and Corley (2002). For economics, Great Britain provides an example: the share of heterodox, not strictly neoclassical economics sank drastically since the ranking of departments became based mainly on citation counts. Heterodox journals have become less attractive for researchers due to their smaller impact factor when compared to mainstream journals (Lee 2007; see also Holcombe 2004)

Fourth, the *establishment of new research areas is inhibited*. In Great Britain, the Research Assessment Exercise has discouraged research with uncertain outcomes and has encouraged projects with quick payoffs (Hargreaves Heap 2002).

Fifth, it is argued that a *positional competition* or a rent-seeking game takes place instead of an enhancement of research quality by the increased investment by universities and journals in evaluating research (Ehrenberg 2000). It has been shown that the percentage of "dry holes" (i.e., articles in refereed journal which have never been cited) in economic research during 1974 to 1996 has remained constant (Laband and Tollison 2003), though the resources to improve the screening of papers have risen substantially.

With respect to *motivational aspects* of rankings on the institutional level a negative selection effect is to be expected, in particular, when monetary rewards are linked to the position in rankings. According to Merton (1973), a special incentive system called "taste for science" exists in academia. It is characterized by a relatively low importance of monetary incentives and a high importance of peer recognition and autonomy. People are attracted to research for which, at the margin, the autonomy to satisfy their curiosity and to gain peer recognition is more important than money. They value the possibility of following their own scientific goals more than financial rewards. These scholars are prepared to trade-off autonomy against money, as empirically documented by Stern (2004): scientists pay to be scientists. The preference for autonomy to choose their own goals is important for innovative research in two ways. It leads to a useful self-selection effect, and autonomy is the most important precondition for intrinsic motivation, which in turn is required for creative research (Amabile et al. 1996; Amabile 1998; Mudambi, R., Mudambi, S., and Navarra 2007)

Are there Alternatives to Academic Rankings?

As discussed, academic rankings have advantages and disadvantages. So far, it cannot be decided whether the advantages of rankings outweigh the disadvantages. The intended advantages consist of more transparency and control of research by non-experts as it is expressed by the view of new public management. The disadvantages consist on the one hand in the technical and methodological problems which might be overcome sometime in the future. On the other hand they consist in the behavioral reactions of reactivity and motivation disturbances which remain even if the indicators were perfect. As a consequence, there is the danger that "the very action of controlling universities and making them more accountable leads them to give a less good account" (Hargreaves Heap 2002, p. 388). The question arises whether there is a third way for research governance which makes use of peer reviews and rankings to a certain degree, but limits its importance for academic careers.

To answer this question we refer to insights from managerial control theory (e.g. Thompson 1967; Ouchi 1977, 1979; Eisenhardt 1985, Schreyögg and Steinmann 1987; Simons 1995). According to this approach there exist three types of control systems: output control, process control, and clan control. The type of control applied must fit the knowledge available to the controller (Turner and Makhija 2006) with respect to outcome measurability and process relations.

Output control is useful if well-defined unambiguous indicators are available to the evaluator, while knowledge of cause-effect or process relations is not necessary. Therefore output controls are attractive to non-experts. As we have discussed, rankings are far from delivering unambiguous indicators to non-experts and should therefore be used with utmost care. *Process control* is useful when outputs are not easy to measure and to attribute, but when the controller is knowledgeable on process relations whose correctness is to be evaluated ex post. Therefore process control is applicable only for peers who are familiar with the state of the art about processes and methodologies in the respective research field. As discussed, peer control has many shortcomings and is particularly questionable when unorthodox contributions have to be evaluated. In such cases well established standards of methods often are challenged. If neither output control nor process control work sufficiently then *clan control* has to be applied (Ouchi 1977, 1979). Clan control is defined as a form of input or ex ante control, based on careful selection and socialization. The aim is to make candidates members of a community in which aligned norms and values are internalized and are part of their intrinsic motivation. If input control is successful, mutual tolerance for ambiguity is possible, which is important when output measurement is questionable and

procedural rules are in flux.

What does clan control mean in the case of research governance? Aspiring scholars should be carefully socialized and selected by peers to show that they master the state of the art, have preferences according to the "taste for science" (Merton 1973), and are able to direct themselves. Those passing a rigorous input control should be given much autonomy to foster their creativity and intrinsic motivated curiosity. This includes the provision of basic funds to give a certain degree of independence after having passed the entrance barriers (Gillies 2008; Horrobin 1996).

Clan control still requires to some extent peer evaluations. However, this applies during restricted periods, namely during the selection and socialization process and when scholars apply to a new position or for a grant, or submit a paper to a journal. However, there is a great difference between being under pressure to publish permanently one the one hand, and being submitted to control during a restricted phase on the other hand, knowing that once this phase is over one will enjoy a wide range of autonomy. Moreover, clan control is better able than output control to use different indicators in an informed way taking their weaknesses into account.

Input or clan control was recommended by the famous President of Harvard University James Bryan Conant: „There is only one proved method of assisting the advancement of pure science – that is picking men of genius, backing them heavily, and leaving them to direct themselves" (Renn 2002).[17] This view is still part of the „Principles Governing Research at Harvard", stating: „The primary means for controlling the quality of scholarly activities of this Faculty is through the rigorous academic standards applied in selecting its members."[18]

Such governance principles are also employed in other professions characterized by a low degree of observable outputs, such as in the life-tenured American judiciary (e.g. Benz and Frey 2007; Posner forthcoming). These ideas are in accordance with empirical findings in psychological economics. They show that on average intrinsically motivated people do not shirk when they are given autonomy (Frey 1992; Gneezy and Rustichini 2000; Fong and Tosi 2007). Instead, they raise their efforts when they perceive that they are trusted (Falk and Kosfeld 2006; Osterloh and Frey 2000; Frost, Osterloh, and Weibel forthcoming).

A comparison between two Australian Universities with similar research interests illustrates the usefulness of clan control (Butler 2003). In the late 1980 the University of Western Australia distributed research funds according to publication counts as the main

[17] Letter to the New York Times, 13. August 1945.

[18] See http://www.fas.harvard.edu/research/greybook/principles.html.

criterion. The University of Queensland followed a different strategy, recruiting bright young researchers and providing them with a strong resource base. Both universities succeeded in lifting their publications per researcher. But only the University of Queensland was successful in improving the quality of publications whereas the University of Western Australia fell below the average Australian score.

Clan control has advantages and disadvantages. The *disadvantages* consist first in the danger that some scholars that have passed the selection might misuse their autonomy, reduce their work effort and waste their funds. But this is the price that has to be paid for the potential high performers to flourish. It will be the lower the more rigorous the selection process is conducted. As a consequence, recruiting is by far the most important issue for academic self-governance. Second, clan control is in danger of being submitted to groupthink (Janis 1972). This danger can be overcome by fostering diversity of scholarly approaches within the relevant peer group. The *advantages* consist in downplaying the unfortunate consequences of rankings while inducing young scholars to learn the professional standards of their discipline under the supporting assistance of peers. This support allows to balance the internal tension of scientific work between conformity and originality. "The professional standards of science must impose a framework of discipline and at the same time encourage rebellion against it" (Polanyi 1962/2002, p. 470). Another advantage might consist in the fact that the provision of basic funds to those that have passed the entrance barriers might increase diversity of research approaches (Gillies 2008) and helps to avoid inefficient "research empires" subject to a decreasing marginal effect of additional research resources (Horrobin 1996; Viner, Powell, and Green 2004). While there exists some empirical work in this regard (Etzkowitz and Leydesdorff 2000; Jansen et al. 2007), this issue needs further research.

Conclusion

This paper argues that academic rankings have major disadvantages which tend to be disregarded or downplayed both in the literature and in practice. Rigorous selection and socialization should play a major role in research governance. In contrast, rankings should be attributed lesser importance. This does not mean a return to the old system of "academic oligarchy". Rather, a new balance is sought between "public management" and the "republic of science". This change in academic governance cannot be started and achieved by individual scholars because of the lock-in effect but needs more far-reaching institutional changes. In particular, the bodies overseeing the research system need to take the shortcomings of solely or mainly relying on rankings into account and pay more emphasis on

the selection and socialization process that provides the basis of academic excellence.

References

Abramo, G. D, Angelo, C. A. and Caprasecca. A. 2009. Allocative efficiency in public research funding: Can bibliometrics help? *Research Policy* 38: 206–215.

Adler, R., Ewing, J., and Taylor, P. 2008. *Citation statistics, A report from the joint committee on quantitative assessment of research* (IMU, ICIAM, IMS). A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS).

Adler, N. J. and Harzing, A.-W. 2009. When knowledge wins: Transcending the sense and nonsense of academic rankings. *Academy of Management Learning*, 8: 72–95.

Albers, S. 2009. Misleading rankings of research in business. *German Economic Review,* 3: 352-363.

Amabile, T. 1996. *Creativity in context: Update to the social psychology of creativity*. Boulder.

———. 1998. How to kill creativity. *Harvard Business Review*, 76: 76–87.

Amabile, T. M., Conti, R., Coon, H., Lazenby, J., and Herron, M. 1996. Assessing the work environment for creativity. *Academy of Management Journal*, 39: 1154–1184.

Armstrong, J. S. 1997. Peer review for journals: Evidence on quality control, fairness, and innovation. *Science and Engineering Ethics*, 3: 63–84.

Arrow, K. 1962. Economic welfare and the allocation of resources for invention. In R. Nelson (Ed.), *The rate and direction of inventive activity: Economic and social factors*: 609–626. Princeton, NJ: Princeton University Press.

Bedeian, A. G. 2003. The manuscript review process: The proper roles of authors, referees, and editors. *Journal of Management Inquiry*, 12: 331–338.

———. 2004. Peer review and the social construction of knowledge in the management discipline. *Academy of Management Learning and Education*, 3: 198–216.

Benz, M. and Frey, B. S. 2007. Corporate governance: What can we learn from public governance? *Academy of Management Review*, 32 (1): 92–104.

Bleiklie, I. and Kogan, M 2007. Organization and Governance of Universities. *Higher Education Policy*, 20: 477-493.

Bok, D. 2003. *Universities in the marketplace. The commercialization of higher education*. Princeton and Oxford.

Bornmann, L., Mutz, R., Neuhaus, C., and Daniel H. D. 2008. Citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics,* 8 (June): 93–102.

Brook, R. 2003. Research survival in the age of evaluation. In *Science between evaluation and innovation: A conference on peer review*: 61–66. München: Max-Planck-Gesellschaft.

Browman, H. I. and Stergiou, K. I. 2008. Factors and indices are one thing, deciding who is scholarly, why they are scholarly, and the relative value of their scholarship is something else entirely. *Ethics in Science and Environmental Politics*, 8: 1–3.

Bush, V. 1960 [1945]. *Science: The endless frontier.* Report to the president. Washington DC: National Science Foundation.

Butler, L. 2003. Explaining Australia's increased share of ISI publications—the effects of a funding formula based on publication counts. *Research Policy,* 32: 143–155.

———. 2007. Assessing university research: a plea for a balanced approach. *Science and Public Policy,* 34: 565–574.

Campanario, J. M. 1996. Using citation classics to study the incidence of serendipity in scientific discovery. *Scientometrics,* 37: 3–24.

Campbell, D. T. 1957. Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, 54:297-312.

Campbell, P. 2008. Escape from the impact factor. *Ethics in Science and Environmental Politics,* 8 (June): 5–7.

Cicchetti, D. V. 1991. The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation. *Behavioral and Brain Sciences*, 14: 119–135.

Clark, B. R. 1998. *Creating entrepreneurial universities. Organizational pathways of transformation.* Surrey: Pergamon Press.

Cole, S. 1992. *Making science. Between nature and society.* Cambridge, MA: Harvard University Press.

Dasgupta, P. and David, P. A. 1994. Toward a new economics of science. *Research Policy*, 23: 487–521.

Daily, C. M., Dalton, D. R., and Cannella, A. A. 2003. Corporate Governance: Decades of

Dialogue and Data. *Academy of Management Review*, 28: 371–382.

Deci, E. L., Koestner, R., and Ryan, R. M. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125: 627–668.

Dogan, M. 1999. Marginality. *Encyclopedia of Creativity*, 2: 179–184.

Donoghue, F. 2008. *The last professors. The corporate university and the fate of the humanities.* New York: Fordham University Press.

Donovan, C. 2007. The qualitative future of research evaluation. *Science and Public Policy,* 34: 585–597.

Durso, T. W. 1997. Editor's advice to reject authors: Just try, try again. *The Scientist*, 11: 13.

Earley, P. C., Connolly, T., and Ekegren, G. 1989. Goals, strategy development, and task performance: Some limits on the efficacy of goal setting. *Journal of Applied Psychology*, 74: 24–33.

Ehrenberg, R. G. 2000. *Tuition rising: Why college costs so much*. Cambridge, MA: Harvard University Press.

Eisenberger, R. and Cameron, J. 1996. Detrimental effects of reward: Reality or myth? *American Psychologist* 51 (11): 1153–1166.

Eisenhardt, K. M. 1985. Control: Organizational and economic approaches. *Management Science*, 31 (2): 134–149.

Espeland, W.N. and Sauder, M. 2007. Rankings and Reactivity: How Public measures Recreate Social Worlds. *American Journal of Sociology* 113 (1): 1–40.

Ethiraj, S. K. and Levinthal, D. 2009. Hoping for A to Z while rewarding only A: Complex organizations and multiple goals. *Organization Science*, 20: 4–21.

Etzkowitz, H. and Leydesdorff, L. 2000. The dynamics of innovation: From national systems and "mode 2" to a triple helix of university–industry–government relations. *Research Policy* 29: 109–123.

Falk, A. and Kosfeld, M. 2006. The hidden cost of control. *American Economic Review*, 96: 1611–1630.

Fehr, E. and Schmidt, K. M. 2004. Fairness and incentives in a multi-task principal-agent model. *Scandinavian Journal of Economics*, 106: 453–474.

Ferraro, F., Pfeffer, J., and Sutton, R. I. 2005. Economics language and assumptions: How theories can become self-fulfilling. *Academy of Management Review*, 30: 8–24.

Fong, E. A. and Tosi, H. L., Jr. 2007. Effort, performance, and conscientiousness: An agency theory perspective. *Journal of Management*, 33: 161–179.

Frey, B. S. 1992. Tertium datur: Pricing, regulating and intrinsic motivation. *Kyklos*, 45: 161–185.

———. 1997. *Not just for the money: An economic theory of personal motivation*. Cheltenham, UK.

———. 2003. Publishing as prostitution? – Choosing between one's own ideas and academic success. *Public Choice,* 116: 205–223.

———. 2009. Economists in the PITS. *International Review of Economics*.

Frey, B. S. and Jegen, R. 2001. Motivation crowding theory. *Journal of Economic Surveys*, 15 (5): 589–611.

Frey, B. S. and Neckermann, S. 2008. Awards – A view from psychological economics. *Journal of Psychology,* 216: 198–208.

Frey, B.S. and Rost, K. forthcoming. Do Rankings Reflect Research Quality? *Journal of Applied Economics.*

Frost, J., Osterloh, M., and Weibel, A. forthcoming. Governing Knowledge Work. Transactional and Transformational Solutions. *Organizational Dynamics.*

Gagné, M. and Deci, E. L. 2005. Self–determination theory and work motivation. *Journal of Organizational Behavior*, 26: 331–362.

Gans, J. S. and Shepherd, G. B. 1994. How are the mighty fallen: Rejected classic articles by leading economists. *Journal of Economic Perspectives*, 8: 165–179.

Garfield, E. 1997. Editors are justified in asking authors to cite equivalent references from same Journal. *British Medical Journal*, 314: 1765.

———. 2006. The history and meaning of the journal impact factor. *The Journal of the American Medical Association (JAMA)*, 295 (1): 90–93.

Gerhart, B. and Rynes, S. L. 2003. *Compensation: theory, evidence, and strategic implications*. Sage.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., and Trow, M. (1994). *The New Production of Knowledge. The Dynamics of Science and Research in*

Contemporary Societies. Thousand Oaks, CA: Sage Publications.

Gillies, D. 2005. Hempelian and Kuhnian approaches in the philosophy of medicine: The Semmelweis case. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36: 159–181.

———. 2008. *How should research be organised?* College Publication King's College: London, UK.

Gilliland, S. W. and Landis, R. S. 1992. Quality and quantity goals in a complex decision task: strategies and outcomes. *Journal of Applied Psychology*, 77(5): 672–681.

Gioia, D. A. and Corley, K. G. 2002. Being good versus looking good: Business school rankings and the Circean transformation from substance to image. *Academy of Management Learning and Education*, 1: 107–120.

Gittelman, M. and Kogut, B. 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. *Management Science* 49 (4): 366–382.

Gneezy, U. and Rustichini, A. 2000. Pay enough or don't pay at all. *Quarterly Journal of Economics*, 115: 791–810.

Gottfredson, S. D. 1978. Evaluating psychological research reports: Dimensions, reliability, and correlates of quality judgments. *American Psychologist*, 33: 920–934.

Hamermesh, D. S. 2007. Viewpoint: replication in economics. *Canadian Journal of Economics*, 40(3):715–733.

Handelsblatt 2010, Economist Ranking, http://www.handelsblatt.com/politik/vwl-ranking/

Haney, W. M. 2002. Ensuring failure: How a state's achievement test may be designed to do just that. *Education Week*, (10 July): 56–58.

Hargreaves Heap, S. P. 2002. Making British universities accountable. In Mirowski, P. and Sent, E.-M. (Eds.), *Science Bought and Sold: Essays in the Economics of Science*: 387–411. Chicago, IL: University of Chicago Press.

Heilig, J. V. and Darling-Hammond, L. 2008. Accountability Texas-style: The progress and learning of urban minority students in high-stakes testing context. Educational Evaluation and Policy Analysis, 30(2): 75–110.

Hirsch, J. E. 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America* 102: 16569–16572. http://dx.doi.org/10.1073/pnas.0507655102.

Holcombe, R. G. 2004. The national Research Council ranking of research universities: Its impact on research in economics. *Econ Journal Watch*, 1: 498–514.

Holmstrom, B. P. and Milgrom, P. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization*, 7: 24–52.

Horrobin, D. F. 1996. Peer review of grant applications: a harbinger for mediocrity in clinical research? *Lancet,* 348: 1293–1295.

IDEAS and RePEc 2008. IDEAS Ranking, http://ideas.repec.org/top/

Janis, I. L. 1972. Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes. Boston, MA: Houghton Mifflin.

Jansen, D., Wald, A., Frenke, K., Schmoch, U. and Schubert, T. 2007. Drittmittel als Performanzindikator der wissenschaftlichen Forschung. Zum Einfluss von Rahmenbedingungen auf Forschungsleistung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 59: 125–149.

Judge, T. A., Cable, D. M., Colbert, A. E., and Rynes, S. L. 2007. What causes a management article to be cited – article, author, or journal? *Academy of Management Journal*, 50: 491–506.

Khurana, R. 2007. *From higher aims to hired hands: the social transformation of business schools and the unfulfilled promise of management as a profession*. Princeton, NJ: Princeton University Press.

Kotiaho, J. S., Tomkin, J. L., and Simmons, L. W. 1999. Unfamiliar citations breed mistakes. *Nature*, 400 (July): 307.

Kruglanski, A. W. 1975. The endogenous-exogenous partition in attribution theory. *Psychological Review*, 82: 387–406.

Kuhn, T. 1962. *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

Laband, D. N. and Tollison, R. D. 2003. Dry holes in economic research. *Kyklos,* 56: 161–174.

Lawrence, P. A. 2002. Rank injustice: the misallocation of credit is endemic in science. *Nature*, 415 (21 February): 835–836.

———. 2003. The politics of publication – authors, reviewers, and editors must act to protect the quality of research. *Nature* 422 (20 March): 259–261.

Lazear, E. P. and Shaw, K. L. 2007. Personnel economics: The economist's view of human resources. *Journal of Economic Perspectives*, 21: 91–114.

Lee, F. S. 2007. The research assessment exercise, the state and the dominance of mainstream economics in British universities. *Cambridge Journal of Economics*, 31: 309–325.

Leibniz, G.W. 1931. Relation de l´état présent de la république des lettres. *Sämtliche Schriften und Briefe* (Akademie-Ausgabe), 4. Reihe, Politische Schriften, Band 1. Darmstadt: 569-571.

Lindsey, D. 1991. Precision in the manuscript review process: Hargens and Herting revisited. *Scientometrics*, 22(2): 313–325.

Locke, E. A. and Latham, G. P. 2009. Has goal setting gone wild, or have its attackers abandoned good scholarship? *Academy of Management Perspectives,* 21, 17–23.

Mahoney, M. J. 1977. Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy Research*, 1: 161–175.

Marginson, S. and Considine, M. 2000. *The enterprise university: Power, governance and reinvention in Australia*. Cambridge: Cambridge University Press.

Merton, R. K.1968. The Matthew effect in science. *Science,* 159: 56–63.

———. 1973. *The sociology of science: Theoretical and empirical investigation.* Chicago, IL: University of Chicago Press.

Miner, L. and McDonald, S. 1981. Reliability of peer review. *Journal of the Society of Research Administrators*, 12: 21–25.

Moed, H. F. 2002. The impact factors debate; the ISI's uses and limits. *Nature*, 415: 731–732.

———. 2007. The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. *Science and Public Policy*, 34: 575–583.

Monastersky, R. 2005. The number that's devouring science. *Chronicle of Higher Education*, 52 (8): A12.

Mudambi, R., Mudambi, S., and Navarra, P. 2007. Global innovation in MNCs: The effects of subsidiary self-determination and teamwork. *Journal of Product Innovation Management*, 24: 442–455.

Nelson, R. R. 1959. The simple economics of basic scientific research. *Journal of Political Economy*, 67: 297–306.

———. 2004. The market economy, and the scientific commons. *Research Policy*, 33: 455–471.

———. 2006. Reflections on "The simple economics of basic scientific research": Looking back and looking forward. *Industrial and Corporate Change*, 15: 903–917.

Nichols, S. L., Glass, G. V., and Berliner, D. C. 2006. High–stakes testing and student achievement: Does accountability pressure increase student learning? *Education Policy Analysis Archives*, 14 (1): 1–175.

Ordonez, L. D., Schweitzer, M. E., Galinsky, A. D., and Bazerman, M. H. 2009. Goals gone wild: the systematic side effects of overprescribing goal setting. *Academy of Management Perspectives*, 23: 6–16.

Osterloh, M. and Frey, B. S. 2000. Motivation, knowledge transfer, and organizational forms. *Organization Science*, 11: 538–550.

———. 2009. Are more and better indicators the solution? Comment to William Starbuck. *Scandinavian Journal of Management*, 25: 225-227.

Oswald, A. J. 2007. An examination of the reliability of prestigious scholarly journals: Evidence and implications for decision-makers. *Economica*, 74: 21–31.

Ouchi, W. G. 1977. The relationship between organizational structure and organizational control. *Administrative Science Quarterly*, 22: 95–113.

———. 1979. A conceptual framework for the design of organizational control mechanisms. *Management Science*, 25: 833–848.

Perrin, B. 1998. Effective use and misuse of performance measurement. *American Journal of Evaluation*, 19: 367–379.

Peters, D. and Ceci, S. J. 1982. Peer review practices of psychological journals: The fate of published articles, submitted again. *The Behavioral and Brain Sciences*, 5: 187–195.

Polanyi, M. 1962. The republic of science: Its political and economic theory. *Minerva*, 1: 54–73, reprinted in Polanyi, M. 1969. *From Knowing and Being.* Chicago: University of Chicago Press: 49-72, re-reprinted in Mirowski, Ph. and Sent E. M. 2002. *Science Bought and Sold. Essays in the Economics of Science.* The University of Chicago Press: 465–485B

Posner, R.A. forthcoming. From the new institutional economics to organization economics: Four applications to corporate governance, government agencies, and legal institutions. *Journal of Institutional and Theoretical Economics*.

Prichard, C. and Willmott, H. 1997. Just how managed is the McUniversity? *Organization Studies*, 18: 287–316.

Propper, C. 2006. Are public sector workers motivated by money? *Bristol: CHERE distinguished lecture.*

Renn, J. 2002. *Challenges from the past. Innovative structures for science and the contribution of the history of science.* Paper presented at the Max Planck Forum 5, Innovative structures in basic decision research, Ringberg Symposium, München.

Schimank, U. 2005. New public management and the academic profession: Reflections on the German situation. *Minerva*, 43: 361–376.

Schreyögg, G. and Steinmann, H. 1987. Strategic control: A new perspective. *Academy of Management Review*, 12: 91–103.

Schweitzer, M. E., Ordonez, L., and Douma, B. 2004. Goal setting as a motivator of unethical behavior. *Academy of Management Journal*, 47: 422–432.

Shanghai Jiao Tong University 2007, Academic Ranking of World Universities, http://www.arwu.org/rank/2007/ranking2007

Simkin, M. V. and Roychowdhury, V. P. 2005. Copied citations create renowned papers? *Annals Improb. Res.*, 11 (1): 24–27.

Simon, H. A. 1985. Human nature in politics – the dialogue of psychology with political science. *American Political Science Review*, 79 (2): 293–304.

Simons, R. 1995. Control in an age of empowerment. *Harvard Business Review*, 73(2): 80–88.

Simonton, D. K. 2004. *Creativity in Science. Chance, Logic, Genius, and Zeitgeist.* Cambridge et al.: Cambridge University Press.

Singh, G., Haddad, K. M., and Chow, S. 2007. Are articles in "top" management journals necessarily of higher quality? *Journal of Management Inquiry*, 16: 319–331.

Smith, R. 1997. Journal accused of manipulating impact factor. *British Medical Journal*, 314: 463.

Sokal, A.D. 1996. A physicist experiments with cultural studies. http://www.physics.nyu.edu/faculty/sokal/lingua_franca_v4/lingua_franca_v4.html

Speckbacher, G., Wentges, P., and Bischof, J. 2008. Führung nicht-erwerbswirtschaftlicher Organisationen: Ökonomische Überlegungen und Folgerungen für das Hochschulmanagement. *Betriebswirtschaftliche Forschung und Praxis,* 60: 43-64.

Starbuck, W. H. 2005. How much better are the most prestigious journals? The statistics of academic publication. *Organization Science*, 16: 180–200.

———. 2006. The production of knowledge. The challenge of social science research. Oxford University Press.

———. 2009. The constant causes of never-ending faddishness in the behavioural and social sciences. *Scandinavian Journal of Management*, 25: 225-227.

Staw, B. M. and Boettger, R. D. 1990. Task revision: A neglected form of work performance. *Academy of Management Journal*, 33: 534–559.

Stephan, P. E. 1996. The economics of science. *Journal of Economic Literature*, 34: 1199–1235.

———. 2008. Science and the university: Challenges for future research. *CESifo Economic Studies*, 54: 313–324.

Stern, S. 2004. Do scientists pay to be scientists? *Management Science*, 50: 835–853.

Stokes, D. E. 1997. *Pasteur's Quadrant: Basic Science and Technological Innovation*. Washington, D.C.: Brookings Institution Press.

Strathern, M. 1966. From Improvement to Enhancement: An Anthropological Comment on the Audit Culture. *Cambridge Anthroplogy,* 19: 1-21.

Swanson, E. 2004. Publishing in the majors: A comparison of accounting, finance, management, and marketing. *Contemporary Accounting Research*, 21 (1): 223–225.

Taylor, M., Perakakis, P., and Trachana, V. 2008. The siege of science. *Ethics in Science and Environmental Politics*, 8 (June): 17–40.

Teichler, U. 1988. *Changing Patterns of the Higher Education System. The Experience of Three Decades.* London: Jessica Kingsley.

The Thomson Corporation (2008a), ISI Web of Knowledge Essential Science Indicators, http://esi.isiknowledge.com/home.cgi

The Thomson Corporation (2008b), ISI Web of Knowledge Journal Citation Report, http://admin-apps.isiknowledge.com/JCR/

Thompson, J. D. 1967. *Organizations in action: social science bases of administrative theory*. McGraw-Hill.

Todd, P. A. and Ladle, R. J. 2008. Hidden dangers of a "citation culture." *Ethics in Science and Environmental Politics*, 8 (June): 13–16.

Tsang, E. W. K. and Frey, B. S. 2007. The as-is journal review process: Let authors own their ideas. *Academy of Management Learning and Education*, 6: 128–136.

Turner, K. L. and Makhija, M. V. 2006. The role of organizational controls in managing

knowledge, *Academy of Management Review*, 31: 197–217.

Ultee, M. 1987. The Republic of Letter: Learned Correspondence, 1680-1720. *Seventeenth Century*, 2 (1): 95-112.

Ursprung, H. W. and Zimmer, M. 2006. Who is the "Platz–Hirsch" of the German economics profession? A citation analysis. *Jahrbücher für Nationalökonomie und Statistik*, 227: 187–202.

van Fleet, D., McWilliams, A., and Siegel, D. S. 2000. A theoretical and empirical analysis of journal rankings: the case of formal lists. *Journal of Management*, 26 (5): 839–861.

van Raan, A. F. J. 2005. Fatal attraction: Conceptual and methodological problems in the rankings of universities by bibliometric methods. *Scientometrics*, 62: 133–143.

Viner, N., Powell, P., and Green, R. 2004. Institutionalized biases in the award of research grants: A preliminary analysis revisiting the principle of accumulative advantage. *Research Policy*, 33: 443–454.

Weibel, A., Rost, K., and Osterloh, M. 2009. Pay for performance in the public sector – benefits and (hidden) costs. *Journal of Public Administration Research and Theory*.

Weingart, P. 1997. From "Finalization" to "Mode 2": Old Wine in New Bottles? *Social Science Information,* 36 (4):591–613.

Weingart, P. 2005. Impact of bibliometrics upon the science system: Inadverted consequences? *Scientometrics*, 62: 117–131.

Weller, A. C. 2001. *Editorial peer review. Its strength and weaknesses.* Medford, NJ: American Society for Information Science and Technology.

Wenneras, C. and Wold, A. 1999. Bias in peer review of research proposals in peer reviews in health sciences. In Godlee, F. and Jefferson, T. (Eds.), *Peer review in health sciences*: 79–89. London: BMJ Books.

Worrell, D. 2009. Assessing business scholarship: The difficulties in moving beyond the rigor-relevance paradigm trap. *Academy of Management Learning*, 8: 127–130.

Zuckerman, H. A. 1992. The proliferation of prizes: Nobel complements and Nobel surrogates in the reward system of science. *Theoretical Medicine*, 13: 217–231.