

The visible hand of research performance assessment

Julian Hamann,

Forum Internationale Wissenschaft, University of Bonn

Abstract Far from allowing a governance of universities by the invisible hand of market forces, research performance assessments do not just measure differences in research quality, but yield themselves visible symptoms in terms of a stratification and standardization of disciplines. The article illustrates this with a case study of UK history departments and their assessment by the Research Assessment Exercise (RAE) and the Research Excellence Framework (REF), drawing on data from the three most recent assessments (RAE 2001, RAE 2008, REF 2014). Symptoms of stratification are documented by the distribution of memberships in assessment panels, of research active staff, and of external research grants. Symptoms of a standardization are documented by the publications submitted to the assessments. The main finding is that the RAEs/REF and the selective allocation of funds they inform consecrate and reproduce a disciplinary center that, in contrast to the periphery, is well-endowed with grants and research staff, decides in panels over the quality standards of the field, and publishes a high number of articles in high impact journals. This selectivity is oriented towards previous distributions of resources and a standardized notion of “excellence” rather than research performance.

Keywords Performance assessment · Research Assessment Exercise · Research Excellence Framework · Stratification · Standardization · Marketization

Introduction

With higher education governance increasingly focusing on research output, additional control mechanisms have been introduced to assess research performance. Performance assessments are supposed to create competitive research markets that govern academia with an invisible hand, guiding a selective allocation of scarce resources to the best performers (Slaughter and Leslie 1999; Teixeira et al. 2004). The contribution asks how invisible the hand of performance assessments really is, and in what regard the structures of academic disciplines are very substantially and indeed visibly changed by assessments.

Universities in the United Kingdom (UK) have a long tradition of research assessments. Since New Public Management orchestrated the retreat of a close government regulation, quasi-markets have become the regulative authority on which universities and departments compete for researchers, students, and financial resources (Deem et al. 2008; Brown and Carasso 2013). From 1986 on, the

decisive assessments for the performance oriented allocation of public funds by the funding councils are delivered every three to seven years and grade the quality of research across dozens of fields and 100 institutions. The surveys are conducted by the four UK higher education funding bodies and based on a peer review system organized in subject-specific assessment panels that are charged with “identifying excellence in the rich diversity of research” (REF 2012) fields they cover. The panels evaluate most notably research output (normally publications) from designated research staff that has been selected by the respective departments. Hence, although it is departments that are graded, the grades are ultimately based on the output of researchers working in them. The Research Assessment Exercises (RAEs) and, since 2014, the Research Excellence Framework (REF)¹ have a significance for the public funding -- and, importantly, perception -- of departments that is unique to the UK. The assessments follow a seemingly straightforward principle: “Institutions conducting the best research receive a larger proportion of the available grant so that the infrastructure for the top level of research in the UK is protected and developed.” (RAE 2001e) The assessments are meant to identify where to allocate scarce resources, and to concentrate them in selected institutions. The current paper reveals two consequences of this policy: first, performance assessments stratify disciplinary fields, and crucially not alongside the distinction of the “best research”, but alongside previous allocations of resources that may not directly be linked to actual research performance. Second, research is standardized oriented towards uniform criteria for “excellence” that foil an open evolution of knowledge and may not always fit with the “rich diversity” of research practices of academic cultures.

Effects and defects of performance assessments in science

Research on the effects of the assessments suggests symptoms of stratification and standardization. Studies have revealed that status assignments create new patterns of authority (Whitley et al. 2010), for example between panel members and the colleagues whose research they are judging (Sharp and Coleman 2005; Sayer 2014), between research active personnel and teaching oriented personnel (Elton 2000; Tapper and Salter 2002), or between university management and the departments

¹ Assessments of the evolution of the RAE/REF over the years are abound (Bence and Oppenheim 2005; Martin and Whitley 2010). While the RAE 2001 is above all characterized by a grade inflation and a subsequent much more concentrated funding policy by the HEFCE, the main change in the RAE 2008 was the introduction of research profiles for each department, based on what proportion of its publications was judged to be of national or international quality. The most important novelty of the REF is that “output quality” (now weighed at 65%) is supplemented with “impact” (20%) and “research environment” (15%) (REF 2011).

they intervene into (Henkel 1999; Yokoyama 2006). These changed patterns of dependency illustrate very visible symptoms of a supposedly invisible hand of market forces.

In addition to new strata of managerial accountability, research has also identified symptoms of adaption to the criteria of performance assessments, coined reactivity of rankings (Espeland and Sauder 2007). According adaptations of academic practice have been identified for most western countries (cf. Benner and Sandström 2000 for Sweden; Harman 2005 for Australia; Münch 2008 for Germany; Leišytė and Westerheijden 2014 for the Netherlands). In the UK, assessments yielded a transfer market for particularly productive researchers (Hare 2003; Bence and Oppenheim 2005), they oriented publication practices towards peer reviewed journals (Talib 2001; Moed 2008), and overall changed cultures and values of academic life (Harley 2002; Lucas 2006). In light of these findings, the visible hand of performance assessments makes it actually more difficult to identify research excellence, since departments' adaptations to the assessment criteria -- and panel members' concern to penalize their own field -- produced ever increasing proportions of "excellent" research (Tapper and Salter 2004; Wooding et al. 2015).

While interventions into the development of disciplines through the definition of assessment criteria are either attributed to governments (Geuna and Martin 2003; Hicks 2012) or to disciplinary elites (Lee 2007; Martin and Whitley 2010), symptoms of standardization are documented for a number of disciplines. In law departments, academic work increasingly concentrates on placing articles in a small number of highly ranked journals (Campbell et al. 1999). In economics, heterodox approaches fall a victim to orthodox assessment panels and criteria geared to mainstream journals (Lee et al. 2013). In medieval history, scholars follow symbolic compliance strategies in order to balance their own research agenda and the priorities of the funding bodies (Kehm and Leišytė 2010). In all cases, assessments seem to have the effect that researchers stay mostly within the mainstream areas of their field in order to maintain chances of receiving funding, while specialized or innovative research -- "the rich diversity of research" the REF (2012) speaks of -- can hardly establish itself.

The article investigates the discipline of history as a case study, since journal articles in high impact journals, the currency that is valued most highly on shorter-termed research quasi-markets (Moed 2008), has less weight in a discipline that traditionally values long-term studies and books, and perceives research and teaching as more of a unit (Kehm and Leišytė 2010; Zuccala et al. 2014). A longitudinal examination allows to reveal the effects one of the most advanced performance as-

assessments systems in Europe has on a discipline with very different research and publication standards. The current article contributes to the state of research by demonstrating, first, how performance assessments and the selective allocation of funds they inform do not just measure differences in research quality but consecrate and reproduce a disciplinary center that is established alongside previous allocations of resources rather than the “best research”. Second, the article suggests that the visible hand of the assessments favors research alongside standardized concepts and criteria that may not always fit with the diverse research practices of academic cultures and hinder an open evolution of academic knowledge.

Performance assessments and analytical heuristics of power

The state of research suggests that performance assessments, rather than enabling dynamics of free research markets, yield themselves symptoms of stratification and standardization. This section discusses possible analytical heuristics that help to grasp and understand these symptoms. While stratification by status indeed appears to oppose free markets that are governed by impersonal calculations of functionality and utility (Weber 1978; Burris 2004), classical functionalist perspectives apprehend the hierarchical structuration of social entities as a functional mean of establishing social order (Davis and Moore 1944). Following universally recognized meritocratic criteria, positions at the top of social hierarchies are occupied by the most qualified actors, a performance elite identified by mechanisms of reward and recognition (Cole and Cole 1973; Merton 1973). With regard to science, the assumption of the functionality of vertical structuration is based on at least two conditions: the indicators applied should be able to measure the complex reality of academic work and not be confined to a small section of it, and the process of vertical structuration should not be influenced by the evaluation procedures themselves. Both, however, is the case for performance assessments. First, indicators that allow for valid measurements of the quality of research, let alone capture the broad range of academic performances, still prove to be a challenge (Laudel 2005; Archambault et al. 2006). Second, linking assessments to the selective allocation of funds triggers a reactivity of rankings. Research is then conducted in a way that corresponds the indicators relevant for the distribution of resources (Espeland and Sauder 2007; Hazelkorn 2007). With the orientation of publication practices towards high impact journals, the recruitment of highly cited researchers, and the staggering general increase in “excellent” research output, the state of research suggests this exact effect for the RAEs/REF.

Hence, the paper applies two related analytical heuristics that seem to be equipped with more explanatory power in order to understand the symptoms of performance assessments (Münch and Schäfer 2014). Bourdieu's (1988) field and capital theory sees procedures of resource distribution as mechanisms of the construction of status hierarchies. They create and reproduce a center-periphery structure in disciplinary fields. The competition on quasi-markets that is actually politically intended in order to identify "excellence" is then foiled by the structurally unequal distribution of social, economic, and symbolic resources (Burris 2004; Morgan 2004). An advocate of the functionalist perspective himself, it was Merton (1968) who warned that a capitalization of research achievements would oppose norms like disinterestedness and communism. According to the Matthew effect, the probability of gains in reputation or resources increases exponentially with every previous gain in reputation or resources until these gains reach a diminishing marginal utility. With regard to the RAEs/REF in history, the article investigates the Matthew effect considering the following research question: *Do structurally unequal opportunities of resource accumulation lead to a stratification of history departments that follows previous allocations of resources rather than research performance?*

Analytical heuristics of power can also explain the second symptom of performance assessments that has been suggested by the state of research. The comprehensive nexus of resource allocation and performance assessments can be understood as an academic panopticon that disciplines research. Foucault's (2010) concept of governmentality can provide a heuristic in order to frame analytically how performance assessments are internalized and act as control mechanisms. While assessments provide accountability for public investment that makes institutions auditable (Power 1997; Strathern 1997), the devised indicators tempt researchers to normalize their work and adjust it to criteria that are important for the allocation of resources (Espeland and Sauder 2007; Morrissey 2013). This standardization perverts any forces a research market could develop. The underlying mechanism has been coined as Campbell's Law: "The more any quantitative social indicator is used for social decision making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social process it was intended to monitor." (Campbell 1979: 85) Regarding the effects of the RAEs/REF on history, Campbell's Law raises the following research question: *Does an adaption to the assessment criteria of the RAEs/REF occur if historians can only benefit from the selective allocation of resources as long as they adopt certain publication practices?*

Data on the RAEs/REF

The two research questions that have been developed from the analytical heuristics are investigated based on data related to the last three assessments of UK history departments. Research was evaluated using a scale from 1 to the top grade 5* (RAE 2001), or grades from 1 to the top grade 4 (RAE 2008, REF 2014). Assessing the quality of research, the criterion applied in 2001 was “excellence”, while the assessments in 2008 and 2014 specified quality as “originality”, “significance”, and “rigor” (RAE 2001c; REF 2012). Informed by the resulting rankings of history departments in regard to research output, rank groups were constructed for analytical purposes: The “top 6” and “top 14” represent departments with the “strongest research output”, the “bottom 6” and the “bottom 14” represent departments with the “lowest research output” according to the respective RAE/REF status hierarchies (table 1).

The research output of departments is assessed as research staff submit publications. Nominating research staff, departments can pursue either more inclusive or more selective strategies. The full time equivalent (FTE) research positions serve as a proxy for social capital, which is defined as the aggregate of resources linked to more or less institutionalized membership in a group (Bourdieu 1986). In this sense, the number of research positions indicates the resources available at a department for, for example, research proposals, reviews, or academic networks (table 2).

Publications and other submissions are assessed by subject-specific panels (table 1). Professional associations and stakeholders from economy and society can nominate members, the funding councils assemble the panels on basis of these nominations. The panels for history consist exclusively of UK historians. Membership in assessment panels indicates symbolic capital, defined as an academic authority that allows to consecrate research by judging over legitimate problem definitions and problem solutions (Bourdieu 1988).

Apart from publications and research staff, departments also file the external research funding they received. This includes, for example, public and private funding, funding from the European Union, and the project specific funding from the research councils that complements the allocation of public funds by the funding councils. External funds indicate the allocation of economic capital (table 3).

Departments can submit up to four publications per staff member. These selected publications constitute the core of assessment (table 4). To determine whether articles and books have been published at central publishers or in core journals, the paper draws on a field of reputable publishers and high impact journals. The 19 publishers include:

Ashgate, Cambridge University Press, Clarendon Press, Cornell University Press, C.H. Beck, Elsevier, Harvard University Press, MIT Press, Oxford University Press, Palgrave Macmillan, Polity Press, Princeton University Press, Routledge, Sage, Stanford University Press, University of Chicago Press, Vandenhoeck & Ruprecht, Wiley-Blackwell, Yale University Press²

In addition to publishers, a field of 126 high impact journals was constructed. It consists of the 20 history journals with the highest average Thomson Reuters five year impact factor from 2009-2013, and of all 106 history journals that received the highest rating in both assessments of the European Reference Indices for the Humanities (The Past Speaks 2011).³ Although the problem remains that, even with considerable effort, prestigious publishers and journals cannot reliably be covered by standardized benchmarks alone, this approach seems to sufficiently factor in not only the longer-term epistemic rhythms in history, but also the diversity and differentiation of the discipline, something that has been called for by critical perspectives on journal rankings in the discipline (Times Higher Education 2008; Royal Society 2009).

The stratification and standardization of UK history

The investigation is based on data for UK history according to the three most recent assessments (RAEs 2001, 2008, REF 2014). Regarding the evaluation of research output, there seems to be a considerable consistency of those departments at the top and at the bottom of the hierarchy (table 1). From 2001 to 2014, 13 departments make it into the “top 14” at least two out of three times. Ten departments are in the “bottom 14” two out of three times. Not a single department is relegated from the top into the bottom rank group, only one department is moving up from the bottom to the “top 14” group.

² The list of publishers is based on previous insights into the discipline of history (Hamann 2014; Gengnagel and Hamann 2014), has been inductively expanded and revised during the investigation, and is widely congruent with the results of Zuccala et al. (2014).

³ A complete list of journals can be requested from the author.

Table 1 Results for history in RAEs 2001, 2008, and REF 2014, and respective memberships in history panels

Groups		RAE 2001	RAE 2008	REF 2014
"Top 14"	"Top 6"	Birkbeck	Imperial College	Birmingham*
		Cambridge*†	Essex	York*
		Durham*	Kent	Sheffield
		East Anglia*†	Liverpool	Southampton*
		King's College*	Oxford*††	Hertfordshire*
		SOAS*†	Warwick	King's College*
		Oxford Brookes	Cambridge*†	Warwick
		LSE	UCL†	Oxford**
		Birmingham	Birkbeck	Exeter*
		Essex	Southampton*†	Cambridge*
		Exeter	Hertfordshire†	Manchester*
		Hertfordshire†	LSE	Leeds
		Huddersfield	Sheffield*	St Andrews**
		Hull*	Aberdeen	UCL*
"Bottom 14"		St Martin's	Goldsmiths	Chichester
		St Mary's	Sheffield Hallam*	Newman
		Westminster	Leeds TAS	Chester
		Worcester	Canterbury CC	Westminster
		York*†	Chichester	Central Lancashire
		Glamorgan	Cumbria	Liverpool Hope
		Bath Spa	Westminster	Leeds Trinity
		Bolton	Gloucestershire	Greenwich
	"Bottom 6"	Chester	Liverpool JMs	St Mary's
		Edge Hill	Edge Hill	Bath Spa
		Liverpool Hope	Northumbria	Sunderland
		Middlesex	Newman College	Anglia Ruskin
		Staffordshire	Wales, Newport	Gloucestershire
		Thames Valley	Worcester	Bishop Grosseteste

* Departments have history Panel member in assessment of the same year.

† Departments have history Panel member in the respective following assessment.

Sources: (RAE 2001a, 2001b, 2008a, 2008b; REF 2014a, 2014b), author's presentation and calculation

For functionalist perspectives, the high stability of status hierarchies might be explained with a stable distribution of the "best research" at the top positions of the hierarchy. Performance assessments actually identify these differences in research quality, and the resource allocation informed by them protects excellent research and its visible contributions in high impact journals, while

obsolete contributions at the other end of the spectrum should be less visible and lastly dispensable (Cole and Cole 1973). In contrast to this -- inevitably simplistic -- sketch of functionalist perspectives, an analysis informed by Foucault's (2010) concept of governmentality highlights the disciplining effects assessment criteria have on the production of knowledge. Drawing on Bourdieu's (1988) field and capital theory, stable hierarchies can be explained with structurally unequal opportunities of capital accumulation. In the following, an unequal distribution of resources will be examined for social, symbolic, and economic capital. Subsequently, a standardization of academic knowledge will be examined with publication patterns of the research submitted to assessments.

Stratification

Is it possible to identify structurally unequal opportunities of capital accumulation that correspond with the stable hierarchy of rank groups? The unequal distribution of resources will be examined for research staff, panel memberships, and external research grants.

Research active staff as social capital

Departments considered to have the strongest research output are much better endowed with research staff compared to those ranked at the bottom. Table 2 illustrates that, on average, the "top 14" departments combine 29% of all FTE research positions in the field, while only 4,9% are located at the "bottom 14". This stratification is even more distinct between the "top 6" (11,2%) and "bottom 6" (1,7% of all research positions).

Table 2 Total of FTE research positions according to RAEs 2001, 2008, and REF 2014

RAE 2001			
Groups	Research Positions		
"Top 6"	197,7		
"Top 14"	358,6		
"Bottom 14"	84,1		
"Bottom 6"	33,1		
Total 2001	1.719,9		
RAE 2008			
Groups	Research Positions	2001's Research Positions in 2008	
"Top 6"	211,11	2001's "Top 6" in 2008	239,4
"Top 14"	530,71	2001's "Top 14" in 2008	444,49
"Bottom 14"	95,11	2001's "Bottom 14" in 2008	98,7

“Bottom 6”	28,91	2001’s “Bottom 6” in 2008	27,7
Total 2008	1.760,91		
REF 2014			
Groups	Research Positions	2008’s Research Positions in 2014	
“Top 6”	181,7	2008’s “Top 6” in 2014 ⁴	246,65
“Top 14”	641,61	2008’s “Top 14” in 2014	577,73
“Bottom 14”	79,88	2008’s “Bottom 14” in 2014 ⁵	109,6
“Bottom 6”	27,5	2008’s “Bottom 6” in 2014	44,7
Total 2014	1785,88		

Sources: (RAE 1992, 1996, 2001a, 2008b; REF 2014b), author’s presentation and calculation

In a longitudinal perspective, the right half of table 2 demonstrates that the departments evaluated best could increase their FTE research posts until the subsequent assessment:

- By 2008, the “top” departments of the RAE 2001 increased their research staff considerably (“top 6“ +21%, “top 14“ +24%). This growth is much higher than the total average increase of research staff (+2%).
- The “top” departments of the RAE 2008 increased their research staff substantially by 2014 (“top 6” +17%, “top 14” +9%). Their growth is markedly higher than the total average increase of research staff in this period (+1%).

The above average movements of researchers to the “top” departments indicate that very good assessments contribute to an improved endowment with research posts. However, staff movements are not a zero sum game. The staff development of the “bottom 6” shows that less successful assessments can lead to a growth in research staff as well:

- The “bottom 6“ departments of the RAE 2001 have less research staff in 2008 (-16%). A however slight overall increase in research posts (+2%) is contrasted by an exodus of researchers from those departments with the worst grades.
- The “bottom 6” departments of the RAE 2008 experienced a growth in research staff by 2014 (+55%). This increase is much higher than the total average growth (+1%).

The movements of research staff show that less positive assessments do not by default imply an exodus of research staff. Unsuccessful performances can indeed lead to reduced funds and therefore

⁴ Imperial College London is not included in the REF 2014 assessment. The number of FTE positions has been taken over from 2008.

⁵ Wales Newport and Cumbria have not been included in the REF 2014. Their number of FTE positions has been taken over from 2008.

cuts in research positions, but just as well to departments being more selective about whose work is submitted to the next assessment, or to heavy investments in research staff and a subsequent rise in research positions. However departments might react to low status allocations⁶, the data convey that good performances generally lead to a marked increase of research staff. On the basis of an unequal distribution of FTE research positions rather than research achievements, performance assessments and resource allocations establish a lasting center-periphery structure between departments in terms of social capital.

Panel memberships as symbolic capital

Relating the RAE/REF results to the composition of assessment panels gives insight into the link between departments' representation in panels and their performance assessment. Furthermore, a comparison with their ranking in the respective previous assessments reveals recruitment patterns over time (table 1).

- 17 history departments were represented in the panel of the RAE 2001. Six of them were ranked among the “top 14” in the same assessment. No department that delegated a colleague into the history panel ended up in the “bottom 14”. Of the 17 panel members in 2001, nine were recruited from departments that received the top grades 5 or 5* in the previous RAE 1996. None of the “bottom 14” departments of 1996 was represented in the panel of 2001 (cf. RAE 1996).
- Of the 17 departments that had colleagues in the panel of the RAE 2008, five got into the “top 14” and only department ended up in the “bottom 14”. Of the 17 panel members in 2008, 13 were recruited from departments that received the top grades 5 or 5* in the RAE 2001. None of the “bottom 14” departments from 2001 sent a colleague into the 2008 panel.
- The 25 panel members of the REF 2014 are spread over 20 history departments. Eleven of those departments are among the “top 14”. No department that delegated a colleague into the assessment panel ended up in the “bottom 14”. Of the 24 panel members in 2014, three quarters were sent by departments that were in the top third of the 2008 assessment. None of the “bottom 14” departments of 2008 was represented in the 2014 panel.

⁶ The flow of research staff should always be seen in proportion to absolute research positions. While a 55% increase in research staff for the “bottom 6” of 2008 corresponds to an absolute growth of 15,8 FTE research positions, the 17% increase of research staff at the “top 6” departments in the same period equals an absolute growth of 35,5 FTE research positions. Although the differences in relative staff increase (55% and 17%) may indicate the contrary, the gap between both rank groups still grows in favor of the top rank group.

The data indicate that, first, in all cases, departments represented in panels received grades above the average grade of all history departments. Second, there is a relationship between the performance ascribed to departments and the recruitment for panels of future assessments. This reveals not only the symbolic weight RAE results have, but also the substantial amount of authority over evaluation criteria that is placed in the hands of very few researchers. With view on the panels, Sharp and Coleman (2005) urge to make panel membership much more representative of higher education in general, and Henkel (1999: 120) refers to “[s]mall components of the invisible colleges of the disciplines that have assumed visible forms as RAE panels and at the same time directly fed into statutory resource decisions.” The fact that departments of panel members are evaluated above average, and that departments with the best grades have the best chances to be represented in the panels of the following assessment, discloses a lasting concentration of symbolic capital in the center of the field of history departments. This allocation does not necessarily follow the distribution of the “best performance”.

Research grants as economic capital

Relating the results of the RAEs 2001, 2008, and the REF 2014 to external funding attracted by respective rank groups reveals a close connection between research funds and status positions (table 3). This is because departments also filed their funding data to the RAEs/REF. These data directly influenced the assessments. In effect, the RAE/REF results therefore ennoble an established center that is already successfully attracting external funding, while the economic periphery can also expect worse overall grades in terms of research output. By consecrating this economic center-periphery structure, the RAEs/REF are part of a self-fulfilling prophecy that rewards those departments with the best assessments that have already attracted the most external funding.

Table 3 External funding of departments according to RAEs 2001, 2008, and REF 2014 in £

Groups	RAE 2001 (reference period 1995-2001)		RAE 2008 (reference period 2001-2007)		REF 2014 (reference period 2008-2013)	
	Funding	Funding per FTE	Funding	Funding per FTE	Funding	Funding per FTE
“Top 6”	1.975.987	9.995	14.123.456	66.901	9.719.731	53.493
“Top 14”	6.590.824	18.379	50.891.667	95.894	51.996.187	81.040
“Bottom 14”	1.201.780	14.290	509.442	5.909	1.031.306	12.911
“Bottom 6”	403.363	12.186	93.333	3.792	535.723	19.481

Sources: (RAE 2001d, 2008c), author’s presentation and calculation

Table 3 illustrates the following patterns:

- The center-periphery structure of absolute and relative external funding becomes apparent in 2001, 2008, and 2014. Over time, the gap between economic center and periphery has widened. The “top 6” departments of the RAE 2001 accumulated five times more research grants than the “bottom 6”, the “top 6” of the RAE 2008 attracted over 150 times more funding than the “bottom 6”, and the “top 6” of the REF 2014 still accumulated 18 times more grants than the “bottom 6”.⁷ The departments rewarded by the RAEs were very successful in raising their rate of external research grants over time, while the periphery could not conform this pressure. This escalation does not only apply to absolute funding, but to almost the same extent to relative funding per researcher.
- Relating funds to FTE research positions, departments placed in the bottom groups tend to have the lowest per head rate of external funds. The “top 14” departments with high numbers of research staff at their disposal attract more external funds per researcher. This relation is more pronounced in the RAE 2008 (figure 1), but also perceivable in the REF 2014 (figure 2). In both cases, Cambridge, Oxford, and UCL (and, if applicable, LSE) stand out as departments with especially high numbers of research positions.⁸ At the same time, it is exactly the examples of Oxford and Cambridge with relatively low funding rates per head that demonstrate that it is in fact cumulative magnitude effects rather than research performance guide the allocation of financial resources.

⁷ The extraordinary gap between the top and bottom groups in the RAE 2008 is caused by the financial position of the history department at UCL described in footnote 6.

⁸ The peculiar financial position of the history department at UCL in figure 1 is caused by an institutional exception: from 1966 to 2012, the UCL Centre for the History of Medicine was primarily funded by the Wellcome Trust. Accordingly, during the period in question the trust awarded the Centre two grants, which explain the exceptional position of UCL in terms of external research funds (RAE 2008c).

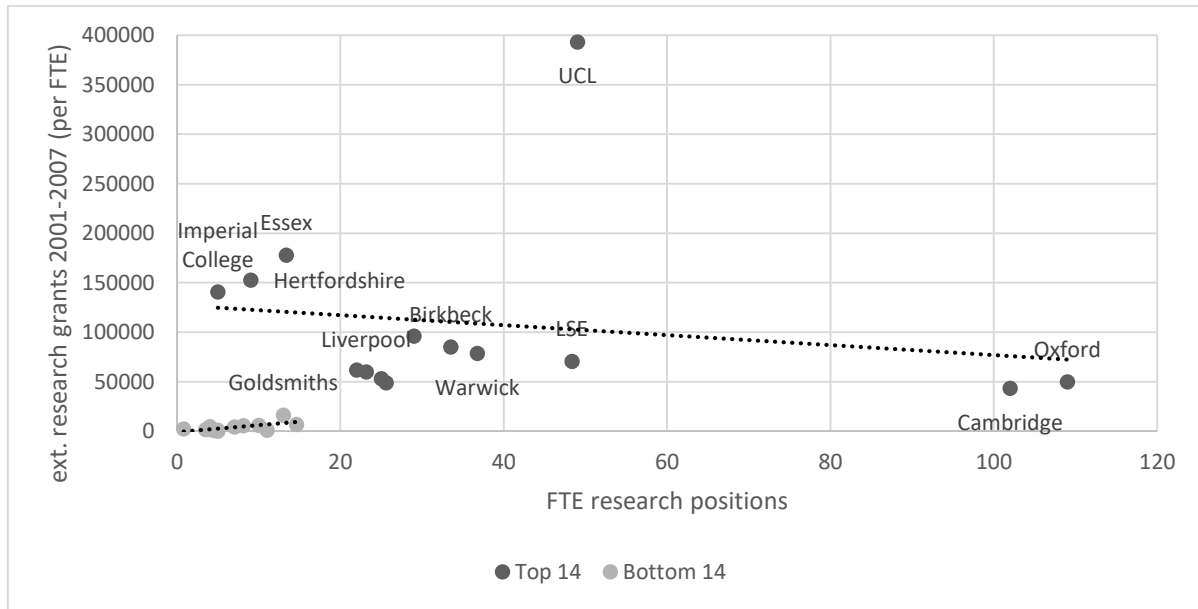


Figure 1 FTE research positions and external research grants 2001-2007 (per research position) for “top 14” and “bottom 14” departments according to the RAE 2008
 Sources: (RAE 2008b, 2008c), author’s presentation and calculation

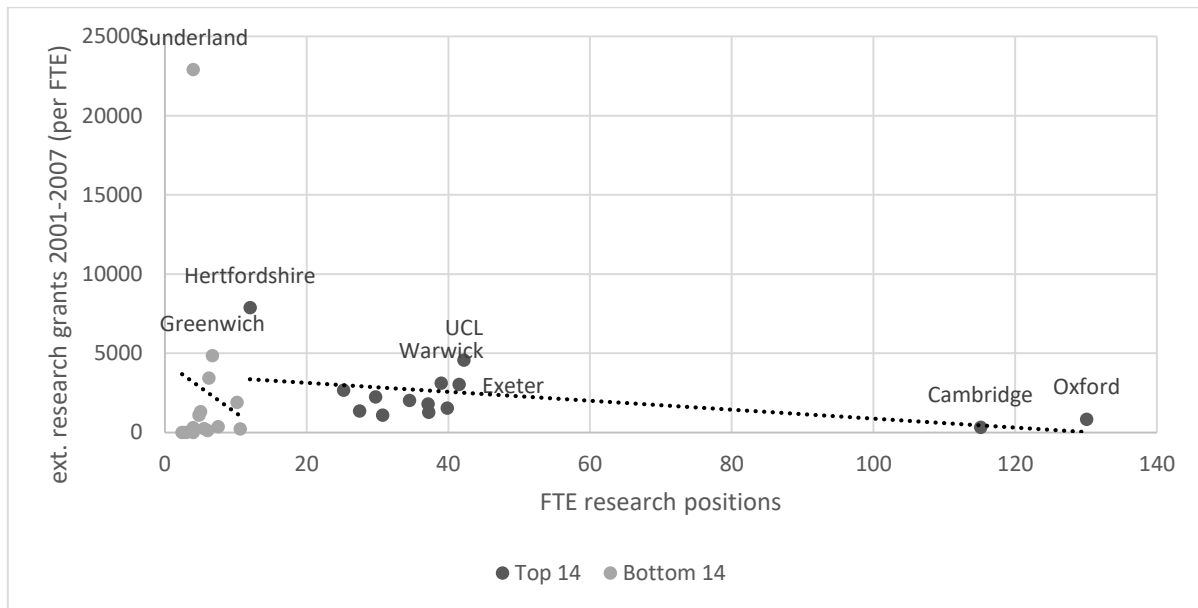


Figure 2 FTE research positions and external research grants 2001-2007 (per research position) for “top 14” and “bottom 14” departments according to the REF 2014
 Sources: (REF 2014b), author’s presentation and calculation

The widening gap between the status groups is hardly surprising considering the mechanism of status allocation: attracting funds is already included as a performance indicator of the assessments and thus directly influences status allocation. Departments that perform well in the assessments are then further endowed with economic capital by the funding councils, and, completing the cycle, it is to be assumed that they also have advantages in future external funding. The status allocations are thus directly involved in the reproduction of an economical center-periphery structure. Relating funding to FTE research positions demonstrates that, crucially, this center-periphery structure does not necessarily correspond with research performance.

Standardization

This section addresses the second research question whether the RAE/REF performance assessments yield a standardization of research and its publication due to an adaption to the disclosed assessment criteria. The examination of the publication profiles of respective RAE/REF status groups has to be sensitive towards the unequal distribution of research positions that has been revealed above. Hence, in a first step, the publication output will be related to the number of research positions of the respective rank groups (table 4).

Table 4 High impact/prestigious publications* submitted to RAE 2008 and REF 2014

Groups		Articles	Monographs
RAE 2008 (publication period 2001-2008)			
"Top 6"	Total	140	117
	Avg. per Person	0,66	0,55
"Top 14"	Total	318	262
	Avg. per Person	0,6	0,49
"Bottom 14"	Total	38	31
	Avg. per Person	0,4	0,33
"Bottom 6"	Total	10	12
	Avg. per Person	0,35	0,42
REF 2014 (publication period 2008-2013)			
"Top 6"	Total	128	78
	Avg. per Person	0,71	0,43
"Top 14"	Total	444	268
	Avg. per Person	0,69	0,42
"Bottom 14"	Total	27	15

	Avg. per Person	0,34	0,19
“Bottom 6”	Total	9	6
	Avg. per Person	0,33	0,22

* Included are all articles in high impact journals, as well as monographs that have been published at prestigious publishers (cf. the section on data).

Source: (RAE 2008c; REF 2014b), author’s presentation and calculation

The data allows the following insights:

- The “top 6“ of 2008 (2014) were able to place 0,66 (0,71) articles per research position in a high impact journal and 0,55 (0,43) monographs at a reputable publisher.

The “top 14“ defined by the RAE 2008 (REF 2014) have published 0,6 (0,69) articles per research position in one of the 126 high impact journals and 0,49 (0,42) monographs at one of the 19 prestigious publishers.

- The “bottom 6“ of 2008 (2014) have published 0,35 (0,33) articles per research position in one of the high impact journals and 0,42 (0,22) monographs.

The “bottom 14“ of the RAE 2008 (REF 2014) could publish 0,4 (0,34) articles per research position in one of the 126 high impact journals and 0,33 (0,19) monographs at prestigious publishers.

There is a close relation between articles per research position that have been published in high impact journals and rank group of departments. A similarly strong relation between publication status and rank group can be found for monographs.⁹ Overall, the differences in publication rates per research position suggest that publications in high impact journals and at prestigious publishers go along with a successful performance in the RAE/REF.¹⁰ Evidently, the “right” publications can be converted into competitive advantages. This normalization might not only standardize how and where research is published. According to the state of research on the standardizing effects of assessment criteria geared to mainstream journals (Willmott 2011; Harley and Lee 1997; Lee and Harley 1998; Lee et al. 2013), it seems likely that there are also effects on research practice and on the content of publications. Future research will have to examine to what extent the narrow corridor of prestigious journals and publishers in history restricts the kind of research that can be rewarded by the RAEs/REF.

⁹ A less pronounced relation could be found between book chapters and rank group, no distinct relation could be found between edited books and rank groups. The data for chapters and edited volumes can be requested from the author.

¹⁰ This is particularly remarkable since the history panel of the REF 2014 vowed not to “privilege any journal or conference rankings/lists, the perceived standing of the publisher or the medium of publication, or where the research output is published.” (REF 2012: 87).

Complementing average publication rates per head for entire status groups, figures 3 and 4 reveal single departments that were outliers in their status groups.

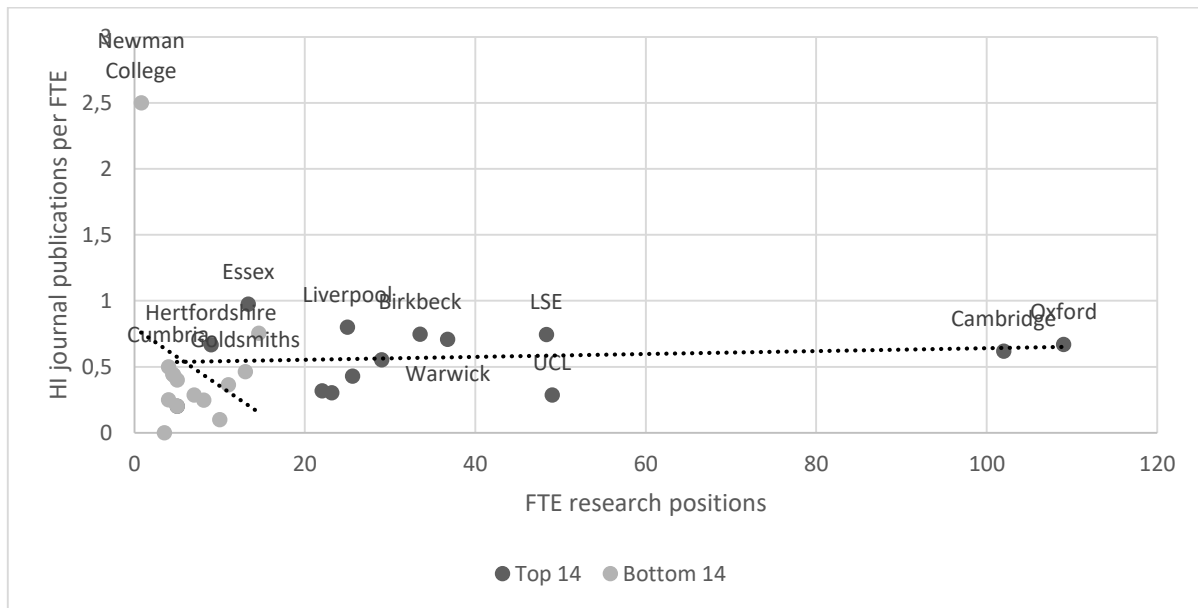


Figure 3 FTE research positions and publications in high impact journals (per research position) for “top 14“ and “bottom 14“ departments according to the RAE 2008

Source: (RAE 2008c), author’s presentation and calculation

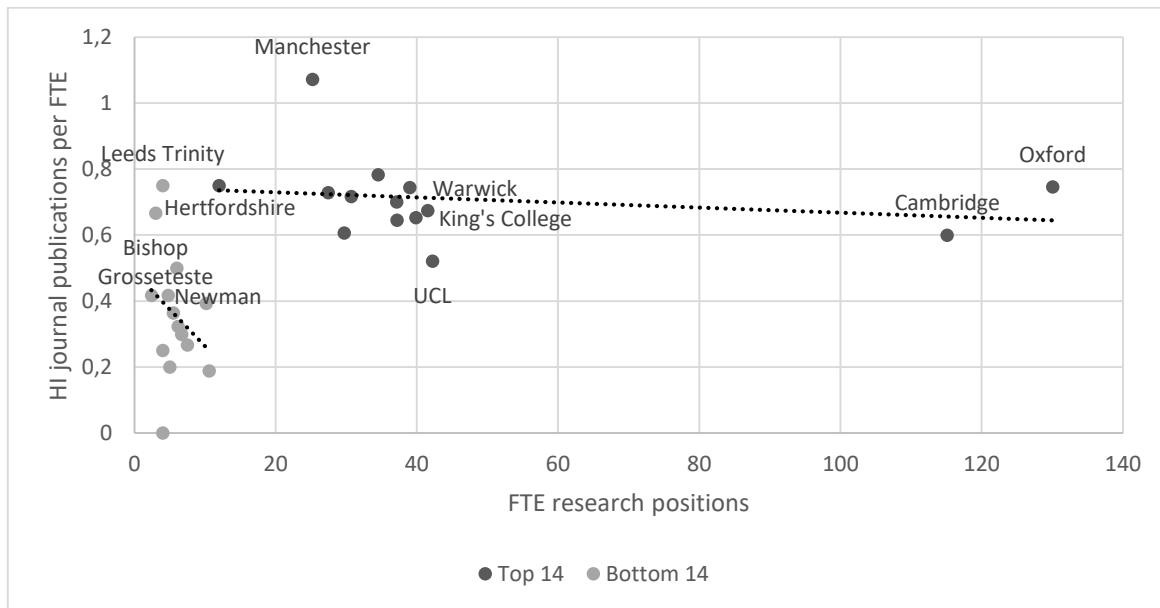


Figure 4 FTE research positions and publications in high impact journals (per research position) for “top 14“ and “bottom 14“ departments according to the REF 2014

Source: (REF 2014b)author’s presentation and calculation

By way of example, the history department at Newman College, located in the “bottom 6” of 2008, could place a relatively high number of articles in high impact journals in relation to a very small number of researchers (figure 3). As has already been the case for external research funding per head, the departments at Oxford and Cambridge cannot confirm their top positions when their publication output is related to the number of researchers.

Although it is instructive to relate publications to the diverging endowment of rank groups with research staff, an exclusive focus on relative publication rates would ignore the magnitude effect the absolute number of publications can develop. Table 4 also illustrates how the differences between rank groups in regard to relative publication rates are reinforced by absolute numbers:

- From the publications the “top 14” departments submitted to the RAE 2008 (REF 2014), 318 (444) were published in high impact journals, while 262 (268) have been monographs at prestigious publishers. The “top 6” departments still placed 140 (128) articles in high impact journals.
- The “bottom 14” departments of 2008 (2014) published only 38 (27) of their submitted publications in high impact journals and 31 (15) monographs at reputable publishers. The “bottom 6” only placed 10 (9) articles in high impact journals.

Due to their mere number of research posts, the “top” departments achieve a visibility that shapes the perception of the field and guides the attribution the “best research”. The longitudinal perspective of table 4 shows that the gap between the “top” and the “bottom” departments in terms of absolute publication numbers has even widened over time. Differentiating the status groups into single departments helps illustrating the magnitude effect (figure 5).

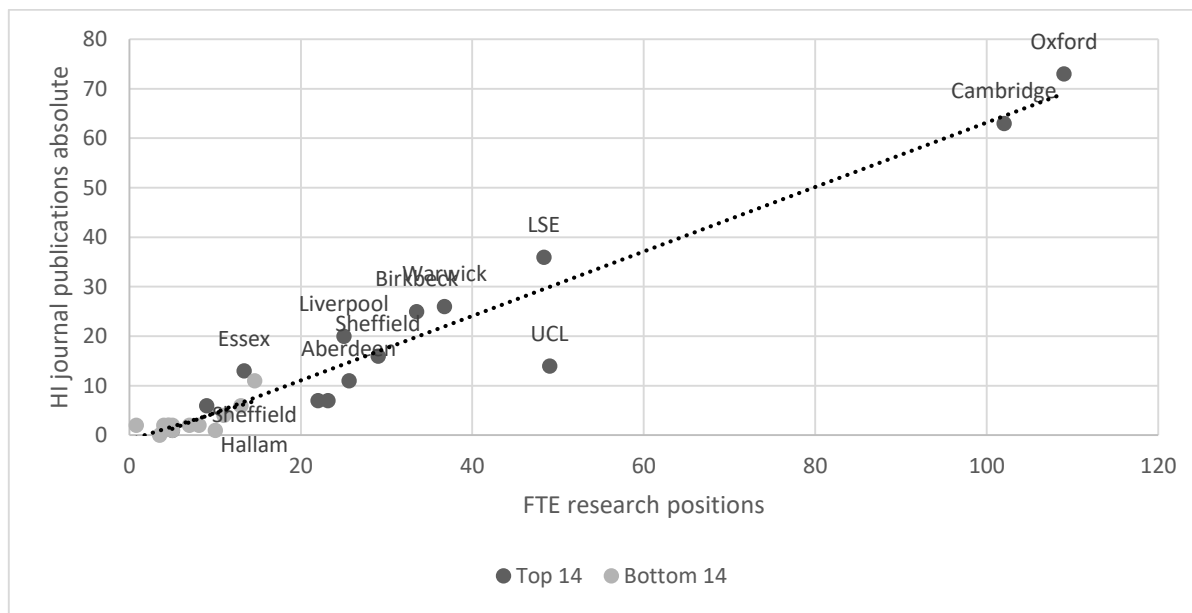


Figure 5 FTE research positions and publications in high impact journals (absolute) from “top 14” and “bottom 14” departments according to the RAE 2008

Source: (RAE 2008c), author’s presentation and calculation

Again, especially Oxford and Cambridge, and also LSE and UCL set themselves apart from the rest of the field with a high amount of articles. This is due to the high number of research posts at their disposal, and despite their relatively low publication rates per head (figures 3 and 4). The plain differences in absolute numbers underline the visibility of the “top” departments and the relative invisibility of the “bottom” ones. From 2001-2008, the “top 14” departments published more than eight times more articles in high impact journals than the “bottom 14”; from 2008-2013, they even published over 16 times more articles in high impact journals.

Departments at the top of the status hierarchy thus constitute a center that is very visible in the field of history departments, in international contexts, and not least for attributions of research performance. Their symbolic weight, coupled with the quantity of their publications, enables them to influence the perception of legitimate publication practices in the disciplinary field as a whole. The RAEs/REF therefore reward not necessarily research performance per head, but the medium of publication, its status, and resources that are already available to departments. The section on stratification has shown that these resources are distributed highly unequal amongst departments and predominantly located in the center of the field.

Discussion

The paper argues that stratification and standardization are symptoms of a visible hand of performance assessment. These symptoms are neither functional nor the kind of selectivity the assessments aim for when they expect to identify and reward “excellence” and the “best research” (RAE 2001e; REF 2012).

The development revealed for history may be less distinct -- or at least not linked to RAEs/REF -- in disciplines that have a proclivity for incremental research within an established paradigm, a weaker link between teaching and research, and a culture of refereeing and linking quality to a hierarchy of journals. It seems plausible to draw this line between laboratory-based sciences on the one hand and social sciences and humanities on the other (Martin and Whitley 2010). Nonetheless, the state of research suggests that the findings apply to varying degrees for most disciplinary fields (Campbell et al. 1999; Kehm and Leišytė 2010; Lee et al. 2013). Since the UK was one of the first countries both to institutionalize research assessment and to link it to financial allocations, the RAEs/REF are one of the most advanced assessment systems in Europe. As suggested by comparative studies on excellence initiatives (Geuna and Martin 2003; Hicks 2012), the symptoms outlined in the current paper are therefore likely to be more profound than in many other states, highlighting developments more clearly that might be less visible elsewhere.

Functionalist perspectives may perceive the stratification as an allocation of scarce resources at the departments with the “best research”. This is led by a belief in the invisible hand of research markets: following a meritocratic logic, high amounts of funding and many publications at reputable publishers and high impact journals are reliable indicators of “excellent” research. The resulting center-periphery structure thus clearly depicts different levels of performance. Hence, members of the “best” departments are in a position to classify the work of their colleagues in assessment panels. The standardization of research and its publications orients the production of knowledge towards expedient standards. Market forces eliminate obsolete contributions and by this objectify knowledge. The functional effect of standardization, therefore, is a necessary reduction of complexities and uncertainties. Assuming a functionalist standpoint, the RAEs/REF can be understood as mechanisms of reward and recognition that help to identify and establish “excellence”.

The empirical evidence suggests that analytical heuristics of power are better equipped to explain symptoms of research assessments. They reveal that assessments, far from an invisible hand meritocratically orchestrating resource allocation, have two visible and lasting effects themselves. First,

performance assessments do follow previous allocations of resources rather than just measuring differences in research quality and rewarding the “best research”. Monopoly profits in terms of social, symbolic, and economic capital are in fact yielded relatively independent from actual research outputs. With rankings, and the allocation of research funds informed by them, the RAEs/REF make a major contribution to producing, reproducing, and consecrating a center-periphery structure. In history, the golden triangle of Cambridge, Oxford, and London-based UCL and LSE is in the very center of the field. Departments with high amounts of research grants and a high number of research staff achieve a high visibility just by the mere number of articles they publish in high impact journals. This is perceived as the “best research”. The longitudinal perspective of the current paper highlighted that, with the resources the “top” departments already have at their disposal and the RAEs/REF “protecting” (RAE 2001e) them, these departments can maintain their position and monopolize reputation. Hence, following the perspective of field and capital theory, the Matthew effect could be confirmed in terms of a stratified distribution of social, symbolic, and economic capital.

In terms of standardization, the RAEs/REF and the selective allocation of resources informed by them foil an open evolution of knowledge by encouraging a conversion of academic knowledge into competitive advantages. Only a very narrow definition of “the best research” is acknowledged: publications in high impact journals and at prestigious publishers. Historians that adapt their publication strategies accordingly by orienting towards visible journals and reputable publishers increase their departments’ chances to retain or improve structural opportunities of capital accumulation. Following the concept of governmentality, the visible hand of performance indicators yields disciplining effects on publication practices, confirming the reactivity of rankings in terms of Campbell’s Law.

With the link between status groups and actual research performance weakened like that, it seems unlikely that an invisible hand could lift the quality of research as is desired by the assessments.

References

- Archambault, É., Vignola Gagné, É., Côté, G., Larivière, V., & Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics*, 68(3), 329-342.
- Bence, V., & Oppenheim, C. (2005). The Evolution of the UK's Research Assessment Exercise: Publications, Performance and Perceptions. *Journal of Educational Administration and History*, 37(2), 137-155.
- Benner, M., & Sandström, U. (2000). Institutionalizing the triple helix: research funding and norms in the academic system. *Research Policy*, 29(2000), 291-301.

- Bourdieu, P. (1986). The Forms of Capital. In J. G. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241-258). New York: Greenwood Press.
- Bourdieu, P. (1988). *Homo Academicus*. Cambridge: Polity Press.
- Brown, R., & Carasso, H. (2013). *Everything for Sale? The Marketisation of UK Higher Education*. London: Routledge.
- Burris, V. (2004). The Academic Caste System: Prestige Hierarchies in PhD Exchange Networks. *American Sociological Review*, 69(2), 239-264.
- Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, 2(1), 67-90.
- Campbell, K., Vick, D. W., Murray, A. D., & Little, G. F. (1999). Journal Publishing, Journal Reputation, and the United Kingdom's Research Assessment Exercise. *Journal of Law and Society*, 26(4), 470-501.
- Cole, J. R., & Cole, S. (1973). *Social Stratification in Science*. Chicago: Chicago University Press.
- Davis, K., & Moore, W. E. (1944). Some Principles of Stratification. *American Sociological Review*, 10(2), 242-249.
- Deem, R., Hillyard, S., & Reed, M. (2008). *Knowledge, Higher Education, and the New Managerialism: The Changing Management of UK Universities*. Oxford: Oxford University Press.
- Elton, L. (2000). The UK Research Assessment Exercise: Unintended Consequences. *Higher Education Quarterly*, 54(3), 274-283.
- Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity. How public measures recreate social worlds. *American Journal of Sociology*, 113(1), 1-40.
- Foucault, M. (2010). *The Government of Self and Others. Lectures at the Collège de France 1982-1983*. New York: Palgrave Macmillan.
- Gengnagel, V., & Hamann, J. (2014). The Making and Persisting of Modern German Humanities. Balancing Acts between Autonomy and Social Relevance. In R. Bod, J. Maat, & T. Weststeijn (Eds.), *The Making of the Humanities III. The Modern Humanities* (pp. 641-654). Amsterdam: Amsterdam University Press.
- Geuna, A., & Martin, B. R. (2003). University Research Evaluation and Funding: An International Comparison. *Minerva*, 41(4), 277-304.
- Hamann, J. (2014). *Die Bildung der Geisteswissenschaften. Zur Genese einer sozialen Konstruktion zwischen Diskurs und Feld*. Konstanz: UVK.
- Hare, P. G. (2003). The United Kingdom's Research Assessment Exercise: Impact on Institutions, Departments, Individuals. *Higher Education Management and Policy*, 15(2), 43-61.
- Harley, S. (2002). The Impact of Research Selectivity on Academic Work and Identity in UK Universities. *Studies in Higher Education*, 27(2), 187-205.
- Harley, S., & Lee, F. S. (1997). Research Selectivity, Managerialism, and the Academic Labor Process: The Future of Nonmainstream Economics in U.K. Universities. *Human Relations*, 50(11), 1427-1460.
- Harman, G. (2005). Australian social scientists and transition to a more commercial university environment. *Higher Education Research & Development*, 24(1), 79-94.
- Hazelkorn, E. (2007). The Impact of League Tables and Ranking Systems on Higher Education Decision Making. *Higher Education Management and Policy*, 19(2), 1-24.
- Henkel, M. (1999). The modernisation of research evaluation: The case of the UK. *Higher Education*, 38, 105-122.
- Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, 41(2), 251-261.
- Kehm, B. M., & Leišytė, L. (2010). Effects of New Governance on Research in the Humanities – The Example of Medieval History. In D. Jansen (Ed.), *Governance and Performance in the German Public Research Sector. Disciplinary Differences* (pp. 73-90). Berlin: Springer.
- Laudel, G. (2005). Is External Research Funding a Valid Indicator for Research Performance? *Research Evaluation*, 14(1), 27-34.
- Lee, F. S. (2007). The Research Assessment Exercise, the State and the Dominance of Mainstream Economics in British Universities. *Cambridge Journal of Economics*, 31(2), 309-325.
- Lee, F. S., & Harley, S. (1998). Peer Review, the Research Assessment Exercise and the Demise of Non-Mainstream Economics. *Capital & Class*, 22(3), 23-51.
- Lee, F. S., Pham, X., & Gu, G. (2013). The UK Research Assessment Exercise and the narrowing of UK economics. *Cambridge Journal of Economics*, 37(4), 693-717.
- Leišytė, L., & Westerheijden, D. (2014). Research Evaluation and Its Implications for Academic Research in the United Kingdom and the Netherlands. *Discussion Papers des Zentrums für Hochschulbildung, Technische Universität Dortmund*, 2014(1), 3-32.
- Lucas, L. (2006). *The Research Game in Academic Life*. Maidenhead: Open University Press.
- Martin, B. R., & Whitley, R. D. (2010). The UK Research Assessment Exercise. A Case of Regulatory Capture? In R. D. Whitley, J. Gläser, & L. Engwall (Eds.), *Reconfiguring Knowledge Production. Changing Authority*

- Relationships in the Sciences and their Consequences for Intellectual Innovation* (pp. 51-80). Oxford: Oxford University Press.
- Merton, R. K. (1968). The Matthew Effect in Science. *Science*, 159(3810), 56-63.
- Merton, R. K. (1973). *The Sociology of Science. Theoretical and Empirical Investigations*. Chicago: University of Chicago Press.
- Moed, H. F. (2008). UK Research Assessment Exercises: Informed judgments on research quality or quantity? *Scientometrics*, 74(1), 153-161.
- Morgan, K. J. (2004). The research assessment exercise in English universities, 2001. *Higher Education*, 48, 461-482.
- Morrissey, J. (2013). Governing the academic subject: Foucault, governmentality and the performing university. *Oxford Review of Education*, 39(6), 797-810.
- Münch, R. (2008). Stratifikation durch Evaluation. Mechanismen der Konstruktion und Reproduktion von Statushierarchien in der Forschung. *Zeitschrift für Soziologie*, 37(1), 60-80.
- Münch, R., & Schäfer, L. O. (2014). Rankings, Diversity and the Power of Renewal in Science. A Comparison between Germany, the UK and the US. *European Journal of Education*, 49(1), 60-76.
- Power, M. (1997). *The Audit Society. Rituals of Verification*. Oxford: Oxford University Press.
- RAE (1992). Universities Funding Council. Research Assessment Exercise 1992: The Outcome. Circular 26/92 Table 62, History. http://www.rae.ac.uk/1992/c26_92t62.html. Accessed 08.08.2015.
- RAE (1996). 1996 Research Assessment Exercise, Unit of Assessment: 59 History. http://www.rae.ac.uk/1996/1_96t59.html. Accessed 08.08.2015.
- RAE (2001a). 2001 Research Assessment Exercise. Unit of Assessment: 59 History. <http://www.rae.ac.uk/2001/results/byuoa/uoa59.htm>. Accessed 08.08.2015.
- RAE (2001b). Panel List History. <http://www.rae.ac.uk/2001/PMembers/Panel159.htm>. Accessed 08.08.2015.
- RAE (2001c). Section III: Panels' Criteria and Working Methods. http://www.rae.ac.uk/2001/pubs/5_99/ByUoA/Crit59.htm. Accessed 08.08.2015.
- RAE (2001d). Submissions, UoA History. <http://www.rae.ac.uk/2001/submissions/Inst.asp?UoA=59>. Accessed 08.08.2015.
- RAE (2001e). What is the RAE 2001? <http://www.rae.ac.uk/2001/AboutUs/>. Accessed 08.08.2015.
- RAE (2008a). RAE 2008 Panels. <http://www.rae.ac.uk/aboutus/panels.asp>. Accessed 08.08.2015.
- RAE (2008b). RAE 2008 quality profiles UOA 62 History. <http://www.rae.ac.uk/results/qualityProfile.aspx?id=62&type=uoa>. Accessed 08.08.2015.
- RAE (2008c). RAE 2008 submissions, UOA 62 History. <http://www.rae.ac.uk/submissions/submissions.aspx?id=62&type=uoa>. Accessed 08.08.2015.
- REF (2011). Assessment framework and guidance on submissions. <http://www.ref.ac.uk/media/ref/content/pub/assessmentframeworkandguidanceonsubmissions/GOS%20including%20addendum.pdf>. Accessed 08.08.2015.
- REF (2012). Panel criteria and working methods, Part 2D: Main Panel D criteria. http://www.ref.ac.uk/media/ref/content/pub/panelcriteriaandworkingmethods/01_12_2D.pdf. Accessed 08.08.2015.
- REF (2014a). Panel Membership, Main Panel D and sub-panels 27-36. <http://www.ref.ac.uk/media/ref/content/expanel/member/Main%20Panel%20D%20membership%2028Sept%202014%29.pdf>. Accessed 08.08.2015.
- REF (2014b). REF 2014 Results & submissions, UOA 30 - History. <http://results.ref.ac.uk/Results/ByUoa/30>. Accessed 08.08.2015.
- Royal Society (2009). Journals under threat: a joint response from history of science, technology and medicine authors. *Qualität in der Wissenschaft*, 34(4), 62-63.
- Sayer, D. (2014). *Rank Hypocrisies. The Insult of the REF*. New York et al.: Sage.
- Sharp, S., & Coleman, S. (2005). Ratings in the Research Assessment Exercise 2001 – the Patterns of University Status and Panel Membership. *Higher Education Quarterly*, 59(2), 153-171.
- Slaughter, S., & Leslie, L. L. (1999). *Academic Capitalism: Politics, Policies, and the Entrepreneurial University*. Baltimore: John Hopkins University Press.
- Strathern, M. (1997). 'Improving ratings': audit in the British University system. *European Review*, 5(3), 305-321.
- Talib, A. A. (2001). The Continuing Behavioural Modification of Academics since the 1992 Research Assessment Exercise. *Higher Education Review*, 33(3), 30-46.
- Tapper, T., & Salter, B. (2002). The External Pressures on the Internal Governance of Universities. *Higher Education Quarterly*, 56(3), 245-256.

These are proofs. The final publication is available at Springer via <http://link.springer.com/article/10.1007/s10734-015-9974-7>,
Higher Education, DOI 10.1007/s10734-015-9974-7

- Tapper, T., & Salter, B. (2004). Governance of Higher Education in Britain: The Significance of the Research Assessment Exercise for the Founding Council Model. *Higher Education Quarterly*, 58(1), 4-30.
- Teixeira, P., Jongbloed, B. W., Dill, D., & Amaral, A. (Eds.). (2004). *Markets in Higher Education. Rhetoric or Reality?* Dordrecht et al.: Springer.
- The Past Speaks (2011). Rankings of History Journals. <http://pastspeaks.com/2011/06/15/erih-rankings-of-history-journals/>. Accessed 08.08.2015.
- Historians decry journal rankings. (2008, 04. Januar 2008).
- Weber, M. (1978). *Economy and Society*, 2 Vol. Berkeley, Los Angeles: University of California Press.
- Whitley, R. D., Gläser, J., & Engwall, L. (Eds.). (2010). *Reconfiguring Knowledge Production. Changing Authority Relationships in the Sciences and their Consequences for Intellectual Innovation*. Oxford: Oxford University Press.
- Willmott, H. (2011). Journal list fetishism and the perversion of scholarship: reactivity and the ABS list. *Organization*, 18(4), 429-442.
- Wooding, S., van Leeuwen, T. N., Parks, S., Kapur, S., & Grant, J. (2015). UK Doubles Its "World-Leading" Research in Life Sciences and Medicine in Six Years: Testing the Claim? *PLoS ONE*, 10(7), e0132990.
- Yokoyama, K. (2006). The effect of the research assessment exercise on organisational culture in English universities: collegiality versus managerialism. *Tertiary Education and Management*, 12(4), 311-322.
- Zuccala, A., Guns, R., Cornacchia, R., & Bod, R. (2014). Can We Rank Scholarly Book Publishers? A Bibliometric Experiment with the Field of History. *Journal of the American Society for Information Science and Technology*, 66(7), 1333-1347.