

Living with the h-index? Metric assemblages in the contemporary academy

Roger Burrows

Abstract

This paper examines the relationship between metrics, markets and affect in the contemporary UK academy. It argues that the emergence of a particular structure of feeling amongst academics in the last few years has been closely associated with the growth and development of 'quantified control'. It examines the functioning of a range of metrics: citations; workload models; transparent costing data; research assessments; teaching quality assessments; and commercial university league tables. It argues that these metrics, and others, although still embedded within an audit culture, increasingly function autonomously as a data assemblage able not just to mimic markets but, increasingly, to enact them. It concludes by posing some questions about the possible implications of this for the future of academic practice.

Keywords: Metrics, higher education, neoliberalism, UK, quantified control, audit culture

Introduction

Something has changed in the UK academy.¹ Many academics are exhausted, stressed, overloaded, suffering from insomnia, feeling anxious, experiencing feelings of shame, aggression, hurt, guilt and 'out-of-placeness' (Gill, 2010).² One can observe it all around; a deep, affective, somatic crisis threatens to overwhelm us (Chow *et al.*, 2010; Evans, 2004; Lynch, 2010; Pelias, 2004; Sparkes, 2007). It is crucial that it does not. We live in a time of exceptional cultural, economic and political crisis and what we are feeling are symptoms of a far broader set of circumstances that demands attention, understanding and, almost certainly, resistance (Boden and Epstein, 2011; De Angelis and Harvie, 2009; Evans, 2004; Gill, 2010; Holmwood, 2010). We know this; yet somehow we feel unable to reassert ourselves. It is crucial that we obtain some sort of critical distance from the specificities of the current conjuncture, and find ways to reaffirm the importance of the use-values that once differentiated our practices from those of others, before the cultural flattening of market economic imperatives subsumes all. As Lock and Martins (2011: 1) express it:

In the bygone world . . . different kinds of institutions embodied various, incommensurable kinds of value. Academic value was not to be identified

with artistic value, nor artistic value with monetary value, and so on. But in our brave new world, it seems that a single final criterion of value is recognized: a quantitative, economic criterion. All else is no more than a means. And there is a single method for ensuring that this criterion is satisfied: quantified control.

This paper aims to make explicit the functioning of such forms of ‘quantified control’ as a means to better understand why we feel as we do. As such it is a small attempt to counter the strong temptation many feel to fall into a state of depressive complicity. We need to better understand ourselves as academics and the conditions under which we now labour; we need to take very seriously all of the magniloquence concerning critical reflexivity that has dominated the social sciences in recent years but which has been largely absent from any consideration about our own contexts, experiences and practices (see though: Gill, 2010; Lynch, 2010; Pelias, 2004; Sparkes, 2007). As part of this it is essential that we come to terms with the changing role that *number* and *numbers* (Adkins and Lury, 2012; Badiou, 2008; De Angelis and Harvie, 2009; Hovland, 2011; Howie, 2005; Kelly and Burrows, 2012) are playing in our contemporary constitution as ‘academics’. What, this paper asks, are the affective consequences of the contemporary ‘metricization’ of the academy?

To understand the contemporary performative functioning of metrics in higher education we need to locate their emergence within broader transformations of university life associated with de-professionalization, proletarianization, the emergence of new academic identities, responses to managerialism, issues of morale and so on (Deem *et al.*, 2007; Watson, 2010; Willmott, 1995; Winter, 1995). However, we also need to locate it within the far broader context of contemporary neoliberalism and, especially, the role of audit cultures within – what we used to think of as – the ‘public sector’.

Neoliberalism, audit culture and beyond

Gane (2012) offers an interpretation of Foucault’s (2008) recently translated lectures delivered at the Collège de France in 1978–1979, that usefully frames much of the discussion that follows. In these lectures Foucault offers a discussion of neoliberalism that breaks with many influential contemporary accounts (Harvey, 2005; Peck, 2008). For Foucault to equate neoliberalism with *laissez-faire* – where the role of the state is largely restricted to supervising the market – is an analytic error. Gane is interested in the manner in which Foucault, by drawing upon the *ordoliberalism*³ of the Freiburg School of political economy, suggests that this relation between state and market under neoliberalism is, in fact, the converse: ‘a state under the supervision of the market rather than a market supervised by the state’ (Foucault, 2008: 116). Under this model, the only mechanism by which the state can legitimate itself is via ‘self-marketization’. For sure, the neoliberal state has to secure the freedom of

markets but it can only do this with authority if it extends the same logic of the market to its own organizational structures and practices. Gane is concerned to decipher what, for Foucault, this logic is. Rather than viewing markets as primarily spaces of *exchange* – as is the case with *laissez-faire* – markets have to be viewed as primarily sites of *competition*. Under his depiction of neoliberalism – as a form of active statecraft within which the state must engage in all manner of ‘internal’ strategies in order to legitimate its power over ‘external’ market processes – it is no longer a matter of *whether* the market impinges upon state activities but *how* it does so. For Gane (2012) this means that privatization strategies are an inherent part of contemporary modes of governmentality and that, as such, where ‘real’ markets cannot be enacted some form of ‘simulated’ market has to be sanctioned. In much of the ‘public sector’ in the UK this has been done through the introduction of audit and the introduction of various forms of metrics that enable systematic comparisons between individuals, departments and institutions. Hitherto, the development of such audit cultures have largely been interpreted as a shift from regimes of *trust* – often tied up with the supposed legitimacy of professional practice – to ones of *accountability* (Holmwood, 2010, 2011; Power, 1997; Strathern, 2000). However, under Foucault’s reading of neoliberal statecraft, this rhetoric of accountability is but a chimera obscuring the real necessity for the state to justify itself to the market by *mimicking* the market (Gane, 2012).⁴

This paper takes this suggestion as a point of departure in order to examine what happens when we reach a point where metric assemblages begin to emerge that are of such ‘complexity’ that they take us to a point ‘beyond the audit culture’; towards a different hegemonic project where systems of ‘quantified control’ begin to possess their own specificity beyond mere auditing procedures; where there develops an ability not just to mimic, but to *enact* competitive market processes. In what follows we attempt to describe the emergence, development, consolidation and, now, ‘autonomization’ of metric assemblages within the UK academy and reflect upon how their performative character contributes to the generation of the structures of feeling that have come to increasingly define contemporary academic life.

The moment of the metrics?

At what point can we begin to date the emergence of the significant metricization of the UK academy? It is the contention of this paper that the marked change in academic sensibilities invoked by new forms of measure only really began to take hold in a significant manner within the last decade or so.⁵ There has been a slow accumulation of layer upon layer of data at various levels, scales and granularities since the mid-1980s, until there came a point of bifurcation; a point at which structures of feeling in the academy were irretrievably altered by the power of these numbers. The ‘bygone world’ identified by Lock and Martins (2011) was the academy of the 1950s through to the early or

perhaps even the mid-1990s. As De Angelis and Harvie (2009: 10) have observed, academic accounts of the life-world of the postwar university, as contained within studies such as Halsey's (1992) *Decline of Donnish Dominion* and even Slaughter and Leslie's (1997) *Academic Capitalism*, concur that 'measure in any systematic form, with accompanying material consequences . . . [is] . . . new. Measure, as we would now recognise it, simply did not exist in the post-war university'.

Confirmation of this could come from any number of different sources⁶ but perhaps the most affecting is that contained within campus fiction. We are fortunate to have a wonderful book-length sociological analysis of this genre. Ian Carter's *Ancient Cultures of Conceit*, published in 1990, examines the postwar campus novel up to 1988 (Carter, 1990: 285). The fictional life-worlds of the academy that it sets out to analyse are ones that many readers will not only recognize but also, perhaps, feel a guilty nostalgia for. The academic world described in *The History Man* (Bradbury, 1975) or *Coming from Behind* (Jacobson, 1983) will be familiar to many older readers. The impact of Thatcherism is brilliantly dissected not only by David Lodge in *Small World* (1984) and *Nice Work* (1988), but also in the volumes by sociologist-cum-novelist Frank Parkin (who sadly died as this paper was being redrafted) such as *Krippendorf's Tribe* (1985) and, especially, *The Mind and Body Shop* (1987). *The Mind and Body Shop* – so amusing at the time – now reads as a pretty standard description of a world where the template of the 'business plan' has become the norm. In the late 1980s one could still construct a comic novel around the idea that the humanities – philosophy no less – could be forced to commercialize its activities and think about its value or 'impact' outside of the narrow domain of the academy. Possibly the later work of Lodge, but certainly the work of Parkin, are really the only ones in Carter's analysis of the oeuvre that begin to touch upon the transformations that were subsequently unleashed upon the academy: a world in which relations between measure and value have become increasingly enacted via code, software and algorithmic forms of power; a world in which the role of number and numbers has come to take political precedence over the aesthetic, the affective and the hermeneutic; and a world in which structures of feeling have been, consequently, fundamentally altered.

Although published almost a decade after the Carter volume, the world of campus fiction caught up with the beginnings of some of these tendencies with the publication of *Overheads* by Ann Oakley (1999) – another sociologist-cum-novelist. Much of this book is concerned with the tussle that the main protagonist, Professor Lydia Malinder, has with colleagues over the introduction of a new mechanism for workload allocation; the development of Teaching and Research Units (TRUs) as a crude attempt to reduce all aspects of the academic labour process to a common metric. Compared to the lived reality of the huge assemblage of various technologies of measurement to which we are now subject, the fictional battle over TRUs now appears somewhat inconsequential. However, it does provide a rough date from which concerns with the

power of metrics are deemed significant enough to warrant comment in the sphere of fiction, and this coincides with the beginnings of an analytic interest within the academy as well (Strathern, 2000). So the ‘moment of the metrics’ – the point of bifurcation – could be thought of as the beginning this century; the point at which academics could no longer avoid the consequences of the developing systems of measure to which they were becoming increasingly subject. Crudely, at some point between the Research Assessment Exercise (RAE) carried out in 1996 and the one conducted in 2001 (Hicks, 2009; Kelly and Burrows, 2012) the moment of the metrics occurred and the structures of feeling that have now come to dominate so much of academic culture began to take hold.

As discussed elsewhere (De Angelis and Harvie, 2009; Kelly and Burrows, 2012) the life-world of the university is now increasingly enacted through ever more complex data assemblages drawing upon all manner of emissions emanating from routine academic practices such as recruiting students, teaching, marking, giving feedback, applying for research funding, publishing and citing the work of others. Some of these emissions are digital by-products of routine transactions (such as journal citations), others have to be collected by means of surveys or other formal data capture techniques (such as the National Student Survey (NSS)) and others still require the formation of a whole expensive bureaucratic edifice designed to assess the ‘quality’ of research (Kelly and Burrows, 2012).

As De Angelis and Harvie (2009) note, different metrics operate at different scales: some at the level of the individual; some at departmental, school or faculty level; some at institutional level; some at national level; and some at international level. However, all are nested or folded into each other to form a complex data assemblage that confronts the individual academic. One could, for example: have an individual h-index (see below) of X ; publish in journals with an average ‘impact factor’ (see below) of Y ; have an undergraduate teaching load below the institutional norm; have a PhD supervision load that is about average; have an annual grant income in the top quartile for the social sciences; work within an academic agglomeration with a 2008 RAE result that places it within the ‘top 5’ nationally; receive module student evaluation scores in the top quartile of a distribution; work within a school with ‘poor’ NSS results, placing it in the bottom quartile for the subject nationally; teach a subject where only Z per cent of graduates are in ‘graduate’ employment 6 months after they graduate, earning an average of just £18,500; work within a higher education institute that is ranked in the ‘top 10’ nationally in various commercially driven ‘league tables’, and within the ‘top 80’ globally, according to others.

It would be quite easy to generate a list of over 100 different (nested) measures to which each individual academic in the UK is now (potentially) subject. However, for our purposes here, we will consider just six domains: citations; workload models; transparent costing data; research assessments; teaching quality assessments; and university league tables. Through the discus-

sion of these we will be concerned to begin to decipher common themes and points of difference in the origins, functioning, effect and affect of these different measures. To be sure, all of our examples were developed at different points in time, by different organizations and for different purposes: citation measures and league tables have been developed by private companies, publishers and newspapers to promote sales and revenue streams; the transparent costing system was developed by Treasury in order to ensure no cross-subsidization between teaching and research; research assessment exercises were initiated by civil servants but then ‘domesticated’ by funding agencies and scholarly bodies and associations; and the measurement of student satisfaction was prompted by politicians concerned to promote a greater level of consumerism amongst students (and their parents) in the higher education system. However, there are commonalities in that all give an emphasis to numeric representation, order and rank, all focus on the ‘measurable’, and all appear to have an interest in promoting competitive changes that alter numbers and ranks over time. The crucial thing though is that together they are all now *experienced* ‘on the ground’ as a more or less ubiquitous melange of measures – full of legacy tensions and contradictions to be sure – that increasingly function as an overarching data assemblage orientated to myriad forms of quantified control; an assemblage that the enactment of which invokes the sorts of affective reactions enumerated at the outset of this paper.

Citation metrics

We might begin by examining what has happened to that most common of academic practices – citing the work of others. Citations indexes were first developed as a tool for the purpose of tracing the history of ideas (Garfield, 1955). The first edition of the Science Citation Index, developed by Garfield, was published in 1961, listing published articles and the number of times they were cited, but it was soon ‘appropriated by users for purposes for which it was unintended – as a tool for evaluating the literature, individuals, institutions, and countries’ (Davis, 2009: 5). It was not long before citation analysis was used to rank academic journals. But it was only with the advent of online commercial databases such as Thomson-Reuters’ *Web of Knowledge*, Elsevier’s *Scopus* and *Google Scholar*, that bibliometric measures really began to emerge within the global knowledge economy as the key measure for what it was that was supposed to count as ‘academic’ knowledge (Larsson, 2009).

There has developed then a ‘citations industry’ within which a large number of different metrics have vied with each other to best encapsulate ‘the popularity, impact and prominence of specific publications, authors and journals’ (Lindgren, 2011: 8).⁷ There is perhaps nothing much wrong with the superficial logic of this.

Researchers . . . publish their findings . . . in academic journals with a peer-review system that serves as a quality assurance. Published works will

stimulate reactions . . . and more important work will trigger more . . . citations . . . than less important works. (Lindgren, 2011: 8)

However, there is now a large body of research literature that queries the

construct validity of citation count as an indicator of scholarly impact . . . researchers have complex citing motives, of which many may be characterized as personal, self-serving or even political rather than professional or scholarship-serving. (Lindgren, 2011: 8)

One might usefully reflect on one's own motives for citing – or not citing – the work of others in a world in which it is no longer the case that such citations just 'provide pellets of . . . peer recognition' (Merton, 2000: 438) but also possess the potential to feed into metrics that, via ever more complex routes, can be monetized. Two numbers in particular have come to the fore in recent years that encapsulate our changing relationship with writing and publishing: the 'h-index' and the 'impact factor'.

The h-index, introduced only a few years ago by Hirsch (2005), has had many articles written about it (see Egghe, 2009). The formulation of the measure is superficially very simple; Hirsch (2005: 16569) proposes that the h-index be 'defined as the number of papers with citation number $\geq h$ ', and suggests that this might provide 'a useful index to characterize the scientific output of a researcher'. The terseness of this definition obfuscates the actual complexity of the measure. If we rank the papers of any given author in decreasing order of the number of citations each paper has generated then the h-index is the highest rank, h , such that the papers on ranks 1, 2, 3, 4, . . . , h each have h or more citations. The actual h-index will depend upon the database of citations interrogated of course. In general h-indices derived from Thomson-Reuters' *Web of Knowledge* and Elsevier's *Scopus* tend to be lower than those derived from *Google Scholar* because coverage in the former two are currently restricted to selected journals whilst coverage by *Google Scholar* is far more inclusive (Bar-Ilan, 2008).⁸

The appeal of this metric is that it combines the number of papers – as an aspect of quantity – and citations – as a supposed aspect of quality, into a single number that is not influenced by 'outliers' at either the top or the bottom of any citations distribution (Egghe, 2009). Now, one could raise any number of objections to the veracity of this metric as a measure of relative academic worth, but that is not the point. Like many of the metrics we will discuss here, it is not the conceptualization, reliability, validity or any other set of methodological concerns that really matter. The index has become reified; it has taken on a life of its own; a number that has become a rhetorical device with which the neoliberal academy has come to enact 'academic value'.⁹ The number is used: to inform the short-listing of candidates for new posts; as an academic 'marketing device' on CVs; as a 'bargaining chip' in professorial salary negotiations; as a variable in statistical models

designed to predict RAE outcomes; to rank colleagues in REF 'preparedness' exercises; in decisions about institutional restructuring; and to inform decisions about whether or not to accept papers written by particular authors in journals.¹⁰

If the h-index functions to instantiate academic 'value' in the individual researcher or research group, then the various metrics that claim to measure the impact that particular academic journals possess provides another mathematical basis for making ever more subtle distinctions within this hierarchy of supposed worth. For many academics where one 'places' articles is now a matter of much deliberation, informed by a myriad of different metrics. The aspiration to get something published in a 'top journal' is thought to be important not just for the individual standing of particular academics, but also because – although journal impact metrics *are not* normally used in formal deliberations within the RAE/REF – there is evidence to suggest that the proportion of outputs published in such 'top journals' is a significant 'shadow metric' 'explaining' RAE outcomes (Kelly and Burrows, 2012). Being able to positively contribute to the RAE/REF – about which more below – is, of course, increasingly viewed as a necessity for appointment and career progression and, as such, is a major factor in the restructuring of the academic labour market. So the ability to identify what the 'top journals' are is no small matter, not just for academics and the institutions in which they labour, but also for the commercial publishing industry within which journal impact factors often function as a proxy for journal viability and, indeed, profitability.

Even before impact metrics can be calculated for journals they first have to be 'accepted' for inclusion in relevant commercial databases. Perhaps the most influential of these is currently the Thompson-Reuters *Journal Citation Reports* (JCR)®. The process by which journals are 'accepted' is rather amorphous but does involve a lengthy application process. By no means are all journals 'accepted', nor do all journals apply for inclusion.¹¹ Once a journal is included in the database a number of competing metrics can be calculated. Perhaps the best-known measure is the Impact Factor (IF), which is a ratio of citations and recent citable items published.¹² Crudely, an IF of 1.0 means that, on average, an article published in the journal is cited just once. An IF of 2 means that, on average an article published is cited twice.

Concern about the 'impact' of 'impact factors' on the sciences has long been voiced, and it is now apparent that similar anxieties exist within the social sciences as well. The IF, like so many of the metrics discussed here, has become grossly reified and it seems that many academics, academic managers, journal editors, publishers and librarians cannot help but reorient their actions towards it. The co-construction of statistical metrics and social practices within the academy (De Angelis and Harvie, 2009; Kelly and Burrows, 2012) is perhaps seen nowhere clearer than in the insertion of these two small numbers – the h-index and the IF – into the mundane realities of academic life.¹³

Workload models and beyond

These measures – whether used explicitly or not in day-to-day management – are just two amongst a great many others that influence the proximate division of labour within the contemporary academy. There are certainly some departments where such measures of ‘research quality’ are rewarded and factored in to workload allocations. More likely, however, is the use of such metrics as more amorphous rhetorical devices deployed by both academics and their managers, in ‘negotiations’ over teaching and administrative loads. Academic labour has always had to be ‘organized’ in some way or another, but with the intensification of work practices that massification and marketization has necessitated has also come the emergence of ever more complex workload models. As De Angelis and Harvie (2009: 10) note, when class sizes and contact hours were lower, when the time-pressures associated with audit and quantified control were all but absent, when the demands to raise ever more external grant income were less, when concerns with ‘strategy’, ‘branding’ and ‘markets’ were hardly ever considered, and when the stresses of publishing were not as intensively driven by concerns over citation metrics, the RAE/REF and so on, then collegiate discussion, the occasional argument and a few rough calculations by someone in authority was normally all that was involved in working out ‘who did what’.

Oakley’s (1999) fictional account of the implementation of an academic workload model based on TRUs culturally registers the beginning of what can now be a computationally highly complex undertaking. Each possible element of the work to be undertaken has to be: conceptualized and operationalized; data has to be collected and verified; and ‘weights’ that need to be attached to each element have to be agreed. The elements that go into the mix vary dramatically of course, but that is the point, the aim of all of this computation is clear enough: to link measure and value; to make heterogeneous activities commensurable; to put everything on the same abstract scale of calculation; to make the ‘immaterial’ labour of intellectual work visible and controllable. As De Angelis and Harvie (2009: 3) so ably express it:

Armies of economists, statisticians, management-scientists, information-specialists, accountants and others are engaged in a struggle to connect heterogeneous concrete human activities on the basis of equal quantities of human labour in the abstract – that is, to link work and capitalist value . . . [there is now an] . . . attempt to (re)impose the ‘law of value’ through . . . measuring . . . immaterial labour. . . [in] higher education in the UK. . . socially-necessary labour-times of ‘immaterial doings’ are emerging and being driven down at the same time as heterogeneous concrete activities are being made commensurable.

The types of ‘immaterial doings’ that are finding themselves represented on modelling spreadsheets are diverse but might include measures such as:

number of lectures over the year (differentiated by undergraduate and post-graduate levels); number of workshops/seminars over the year; number of supervisees; number of dissertations supervised; number of PhD students supervised; number of post-docs and/or research assistants supervised; research grant income; scope of university committee membership; 'allowances' for research 'buy-out'; and 'allowances' for agreed 'major' and 'minor' administrative responsibilities. These various metrics might themselves then have to be 'adjusted' to account for colleagues: who are 'early-career' and therefore might be expected to deliver an overall workload below the norm; who are 'research stars' (on measures such as h-indices and grant income) and thus, *ceteris paribus*, might be expected to take on less teaching and spend more time on research; and, who are 'poor teachers' and who might be asked to take on more non-teaching responsibilities in order to avoid impacting upon other metrics such as the NSS (see below). When such proximate individualized data used in local workload modelling is supplemented – as is happening – by other metrics derived from commercially procured research surveillance systems¹⁴ giving information on matters such as the number and success rates of grant applications, citation metrics, research 'impact' measures and so on, then we have a real-time post-panoptic academic e-surveillance system emerging of significant power and scope.

TRAC and FEC

Overlaying this often very detailed set of local metrics that attempt to capture the academic labour process is another that attempts to do much the same thing, but for very different purposes. The collection of *Transparent Approach to Costing* (TRAC) data only impinges upon the academic life-world periodically, but its nature is such that, although it does not possess the same affective qualities of some of the other metrics enumerated here, it nevertheless never fails to invoke feelings of despondency and frustration; it is a truly hopeless measure, but not only do academics know this, so do their managers and so do most of the government outside of the Treasury. So, it is as significant as it is for our concerns here, not because of its positive veracity and functioning as a metric but because it is emblematic of our own knowing collusion with a joyless bureaucratic auditing process. It is useless, we all know it is useless, and yet still we have to find the time and the energy to 'make it up'.

The whole bureaucratic paraphernalia associated with TRAC originated in 1998 when it was decided that a transparency review of the funding of universities was required in order to improve the 'accountability' for the use of public funds. As part of this it was decided that data was required in order to ascertain the time academics dedicate to various activities – the 'immaterial doings' that we now spend so much time feeding into workload allocation spreadsheets. But the time has to be allocated to a set of categories that might mean something to someone in the Treasury but which bears very little relationship

to the vernacular of concrete academic work. Every few months one is sampled and asked to fill in a 'time-sheet' showing how many hours one has spent on 20 different categories of activity. Because the TRAC has been designed to discover the relative *proportion* of time actually spent on various activities no account is made of the actual number of hours expended. Most colleagues just roughly make the data up, or just send in the same pre-prepared sheet each time, or are tempted to declare themselves to have been on leave all week. The sad and depressing thing is, of course, that this 'dirty data' is then used to 'model' the supposed full economic costs (FEC) of our activities, which in turn inform the costing of grants and other of our activities. We are 'TRACked and FECKed' as Holmwood (2011) would have it; we live on a lie, we know it, and we feel it.

Measuring research 'quality'

University research in the UK has long been subject to something called a 'dual support system' made up of two parts: 'block-grants', in order to provide a general underpinning for research capacity; and funds for specific research grants, made available by competition, administered by research councils.¹⁵ Although it has long been clear on what basis specific research grants are awarded – peer-review and competition – the allocation of block grants was, for a long time, an altogether different matter. Up until the mid-1980s the allocation was opaque. A first, half-hearted, attempt was made in 1985–86 to make the basis for the allocation more explicit. In an era before the full turn to metrics the criteria by which the quality of research was judged in this first Research Selectivity Exercise (RSE), as it was then called, was actually hardly any more transparent than had been existing allocation mechanisms. Each subject area was asked to produce a brief 'research profile', of no more than three sides of A4, within which, it was suggested, might be information on: financial support for research; staff and research student numbers; any measures of research performance deemed significant; a statement of current and likely future research priorities; and the titles of no more than five books or articles produced since 1980 (Bence and Oppenheim, 2005). The 'results' of this first RSE were published in 1986 on what was essentially a four-point scale: 'outstanding', 'better than average', 'about average', or 'below average'.

The second RSE in 1989 was taken far more seriously as it became apparent that the results would significantly impact upon funding allocations. The 1989 RSE was based on 'informed peer review' from 70 advisory groups and panels. This time the panels were provided with far more structured data on research performance including: the number of publications in relation to the number of full-time academic staff; bibliographical details of up to two publications for each full-time member of academic staff; the number and value of research grants and research contracts; and the number of research studentships. This information was used in order to rate each Unit of Assessment (UoA) on a

five-point scale. By the time of the third exercise in 1992 the university sector had been expanded to include the ex-polytechnics. Research for assessment was divided into 72 UoAs. The data collected became more extensive; in addition to each academic nominating two publications, quantitative information on all publications was also required. Each submission was again ranked on a five-point scale. In the fourth exercise in 1996 'research active' staff in 69 UoAs had to provide details of up to four publications and the rating scale was further finessed to become, essentially, a seven-point scale. Further recalibrations in the process were undertaken for the fifth exercise in 2001. The essence of the assessment remained intact but the descriptions attached to the points on the rating scales were reworded somewhat and the scale itself was extended to include a new eighth notch.

For the sixth and most recent exercise, in 2008, the process of research quality assessment was altered fundamentally in order to produce a rating system that could better approximate an interval level of measurement. Rather than each submission being awarded an overall rating, each was given a 'quality profile'. In the case of sociology a panel of 16 peers examined 39 detailed submissions containing information on: four publications for each member of staff submitted; a detailed narrative and statistical data on the research environment; and a narrative on various esteem measures. Once 'released' the quality profiles develop an autonomous 'public life'. They are not only used to inform funding allocation decisions. They are also inserted in other secondary algorithms used to determine eligibility for other types of funding. They are also used as a rhetorical device by institutions in their prospectuses, on their websites and in funding bids. They are also used – along with other metrics – to construct university 'league tables' of various sorts (see below).

The performative power of the RAE/REF has grown in significance since the mid-1980s. Throughout this period not only has the information gathering involved in the various exercises become ever more detailed and prescribed but there has also been an aspiration to shift the level of measurement used to assess the quality of research from an ever more differentiated ordinal scale towards one closer to a cardinal level of measurement. Increasingly many of the mundane realities of academic life have been recursively lived through not only the exercises themselves, but also various institutional imaginings of what such future exercises might bring.

Measuring teaching 'quality'

The quality of teaching has long been evaluated internally to institutions. There was a time when this involved little more than students completing a short paper-based survey at the end of a course for the consumption and then action of the tutor(s) involved. Such processes have now often become computerized and all manner of measures – quantitative and qualitative – are

available to managers to compare and contrast the feedback students provide to different members of staff. Staff themselves are encouraged to use such feedback in promotion applications and managers are asked to use the data produced to rank staff when providing references; more often than not such rankings are also freely available to colleagues as they are often discussed as part of auditing and reviewing procedures. All of this contributes towards many of the feelings expressed by Gill (2010) at the beginning of this paper. Not only are we being publically judged and ranked but we are also being encouraged to reorient our pedagogical practices towards the preferences, tastes and mores of an ever more consumerist student audience; a recipe for further dissatisfaction on the part of academics as they often struggle with new technologies and new forms of 'infotainment' in the vain hope that they might 'improve' their scores. This might matter less than it does, but such audience responses also feed through to metrics that are published nationally – such as the NSS – which are explicitly comparative and inform the construction of various other institutional and subject 'league tables' constructed from a variable hodgepodge of such publically available measures.

Since 2005 universities have had to participate in the NSS; part of a quality assurance framework integral to the audit culture. The stated aim of the NSS is to gather feedback on the quality of courses in order to contribute to public accountability as well as to help inform the choices of future applicants to higher education. From 2012 some of the NSS will also be available via, what are to be called, Key Information Sets (KIS), which aims to provide ever more detailed information to prospective students, their parents and their advisers. The KIS will contain up to 20 different data items, including eight questions from the NSS, but also data on contact hours; the mix of assessment methods; costs, fees and financial support; accommodation costs; employment and salary information, such as the destinations of graduates six months after completing their course and, of those in employment, the proportion in managerial/professional jobs; salary data after both 6 months and 40 months after graduation; and so on.¹⁶

League tables

Just as at an individual level workload planning attempts to 'flatten' the academic labour process to a common set of metrics, then university league tables do much the same thing at an institutional level. Commercial interests across the globe – mainly publishing concerns – compete to rank institutions, and the subjects within them, based upon their own scales calculated by combining any number of different metrics (only some of which have been discussed here). The degree of variability between the scales suggests that the reliability and the validity of such exercises could be called into question; but, despite this, institutions still make much of their 'position' in their marketing materials and students do appear to draw upon the data in making decisions

about where to apply and so on (Hazelkorn, 2011). Some, such as *Times Higher Education* (THE), have invested heavily in such exercises at both a national and a global level, working in partnership with citation data providers Thomson Reuters, producing websites and even an iPhone ‘app’ that allow users to construct their own bespoke rankings by differentially weighting the various elements that contribute to the overall scale.

In many ways these ‘league tables’ epitomize many of the themes discussed here: they are the result of a whole range of other metrics generated at different levels of individual and organizational life that all become folded and nested into a common scale; they attempt to collapse heterogeneous concrete activities into supposedly commensurable value scales, allowing comparison and competition; they are, themselves, a source of commercial value – providing some sort of a shadow metric of the underlying abstract value of the neoliberal university; and, for the individual academic, they are also inescapable – to work in the academy today inevitably involves enacting intellectual life through such metrics with all of the affective consequences that follow from this.

Conclusion

This paper has examined the relationship between metrics, markets and affect in the contemporary UK academy. It has argued that the emergence of a particular structure of feeling amongst academics in the last few years has been closely associated with the ‘autonomization’ of metric assemblages from auditing processes in the academy. Whereas once metrics were simply part of auditing process and, as such, functioned to ensure *accountability* they have, in more recent times, taken on another role, and now function as part of a process of, what has been called ‘quantified control’. In essence academic metric assemblages are at the cusp of being transformed from a set of measures able to mimic market processes to ones that are able to *enact* market processes. In the neoliberal university world of student fees and ever-greater competition for student numbers and research grant income, these metrics function as a form of measure able to translate different forms of value. Academic *value* is, essentially, becoming monetized, and as this happens academic *values* are becoming transformed. This is the source of our discomfort. It is not just that we might have some political objections to these value transformations. The root of the issue is that – as we have learnt from colleagues working in science and technology studies on the social life methods (Law and Urry, 2004; Saetnan *et al.*, 2011) – we are fully implicated in their enactment. We are all involved in the co-construction of statistics and organizational life (Saetnan *et al.*, 2011). Throughout this paper we have tried to demonstrate how this happens within an academic setting by rehearsing the various interactions we all have with numbers; how we are forced to use the language of statistical measures whether we want to or not; and how, even when on occasion we may

attempt to resist we know that not playing the ‘numbers game’ will have implications for us and our colleagues: ‘play’ or ‘be played’.

The implications of playing or not playing ‘the game’ are, of course, empirical questions. We end with some questions that demand further investigation. Are metrics really experienced as oppressively as they have been characterized here? And, if so, with what ‘impacts’? Are different metrics experienced differently? Do proximate workload models impinge more upon the psyche than the h-index, for example? What are the particular and cumulative effects on person-value? Does it lead to one developing a greater sense of cynicism and/or distancing from institutions, with little in the way of institutional ‘loyalty’? Does it lead to a sense of permanent failure to live up to the idealized academic performing well on all fronts (Lynch, 2010)? And how are such responses and affects distributed? They are almost certainly highly gendered, but likely also linked with generational, ethnic, class and other differences. How do they vary by discipline, institutional type and/or location within the professional hierarchy? What are the wider effects on the academic division of labour, careers, on the student experience and, crucially, on our scholarship? And other than episodic declarations to ‘KIS my FECing AcSS¹⁷’, or suchlike, how do we resist?

Goldsmiths, University of London

Received 4 August 2011
Finally accepted 30 January 2012

Acknowledgements

Thanks are due to Nick Gane, Ros Gill, Mary Holmes, Aidan Kelly, Anne Kerr, Simon Parker and Paul Wakeling for their comments on earlier drafts. Thanks also to the anonymous referees for their erudite suggestions for improvements, which have been drawn upon extensively.

Notes

- 1 Although what follows has a far wider global applicability, as demonstrated by, *inter alia*, Amsler and Bolsmann (2012), Burawoy (2011), Chow *et al.* (2010), Hazelkorn (2011), Shore (2008) and Wright and Rabo (2010).
- 2 Certainly this is the case in the humanities and the social sciences; the position in the STEM subject is, perhaps, more ambivalent.
- 3 The term derives from *The Ordo Yearbook of Economic and Social Order*, an academic journal founded by a group of German political economists in 1948.
- 4 This interpretation runs counter to that found within much of the sociology of higher education (Giroux, 2002; Regini, 2011). It also gives a different emphasis to those who have seen audit, assessment and metrics in higher education as primarily regulatory and statist dynamics (Burawoy, 2011; Neave, 1998).
- 5 Colleagues working in academic planning are adamant that it was the publication of the (now defunct) Committee of Vice-Chancellors and Principals commissioned *Jarratt Report* in 1985, during the apogee of Thatcherism, that the very initial conditions for the thorough metricization of the academy were set.

- 6 See, for example, excellent histories of the period such as Kogan and Hanney (2000) and Henkel (2000).
- 7 There is now a huge body of technical literature on citation analysis and journals devoted to the subject (eg *Scientometrics*).
- 8 As this paper was being redrafted Google introduced their 'My Citations' system as an adjunct to Scholar allowing h-indices to be calculated very easily.
- 9 H-indices in STEM subjects are higher than those in the social sciences, which are, in turn, higher than those in the humanities. There is also considerable variation *within* the social sciences. The LSE blog on the impact of the social sciences estimates that the highest average h-index in the social sciences is found amongst geographers at 5.04, followed by economists at 4.83, political scientists at 2.46, sociologists at 2.38, and law at 1.25. Amongst sociologists, professors in their sample averaged 3.67, senior lecturers 2.50 and lecturers 1.91 (see <http://blogs.lse.ac.uk/impactofsocialsciences/the-handbook/chapter-3-key-measures-of-academic-influence/>).
- 10 An author with a high h-index might be considered more likely to attract future citations which will, in turn, increase the impact factor of the journal (see below).
- 11 In the 2008 RAE submissions to the Sociology Unit of Assessment of the 847 different journals from which 'outputs' were culled only 309 were included in the JCR at the time (Kelly and Burrows, 2012).
- 12 Formally, as Pringle (2008: 87) explains for 2006 the IF would be 'calculated as follows: A = total cites in 2006. B = 2006 cites to articles published in 2004-5 (a subset of A). C = number of articles . . . published in 2004-5. D = B/C = 2006 impact factor . . . '.
- 13 See the recent special issue *Power and Education*, 3 (1), 2010 on 'The Power of Journal Rankings'.
- 14 Systems with names such as *Enquire*, *Metis* and *Pure* that are forming the informational infrastructure to underpin the ongoing metricization of the academy.
- 15 See Kelly and Burrows (2012) for more details.
- 16 See www.hefce.ac.uk/learning/infohe/kis.htm.
- 17 Academician of the Social Sciences.

References

- Adkins, L. and Lury, C. (eds), (2012), *Measure and Value*, Sociological Review Monograph, Oxford: Wiley-Blackwell.
- Amsler, S. and Bolsmann, C., (2012), 'University ranking as social exclusion', *British Journal of Sociology of Education*, 33 (2): 283–301.
- Badiou, A., (2008), *Number and Numbers*, Cambridge: Polity Press.
- Bar-Ilan, J., (2008), 'Which h-index? – A comparison of WoS, Scopus and Google Scholar', *Scientometrics*, 74 (2): 257–271.
- Bence, V. and 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.
- Boden, R. and Epstein, D., (2011), 'A flat earth society? Imagining academic freedom' *The Sociological Review*, 59 (3): 476–495.
- Bradbury, M., (1975), *The History Man*, London: Picador.
- Burawoy, M., (2011), 'Re-defining the public university: Developing an analytical framework', available at: <http://publicsphere.ssrc.org/burawoy-redefining-the-public-university/> (accessed 3 January 2012).
- Carter, I., (1990), *Ancient Cultures of Conceit: British University Fiction in the Post-War Years*, London: Routledge.
- Chow, Y.F., de Kloet, J. and Leung, H.S.S., (2010), 'Towards an ethics of slowness in an era of academic corporatism', *Espaces Temps.net*, <http://espacestemps.net/document8314.html>.

- Davis, P., (2009), 'Reward or persuasion? The battle to define the meaning of a citation', *Learned Publishing*, 22 (1): 5–11.
- De Angelis, M. and Harvie, D., (2009), '“Cognitive capitalism” and the rat-race: How capital measures immaterial labour in British universities', *Historical Materialism*, 17, 3, 3–30.
- Deem, R., Hillyard, S. and Reed, M., (2007), *Knowledge, Higher Education, and the New Managerialism: The Changing Management of UK Universities*, Oxford: Oxford University Press.
- Egghe, L., (2009), 'The Hirsch-index and related impact measures', *Annual Review in Information Science and Technology*, 44: 65–115.
- Evans, M., (2004), *Killing Thinking: The Death of the Universities*, London: Continuum.
- Foucault, M., (2008), *The Birth of Biopolitics: Lectures at the Collège de France 1978–79*, Basingstoke: Palgrave.
- Gane, N., (2012), *Max Weber and Contemporary Capitalism*, Basingstoke: Palgrave.
- Garfield, E., (1955), 'Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas', *Science*, 122: 108–111.
- Gill, R., (2010), 'Breaking the silence: The hidden injuries of neo-liberal academia', in Flood, R. and Gill, R. (eds), *Secrecy and Silence in the Research Process: Feminist Reflections*, 228–244, London: Routledge.
- Giroux, H., (2002), 'Neoliberalism, corporate culture, and the promise of higher education: The university as a democratic public sphere', *Harvard Educational Review*, 72 (4): 424–463.
- Halsey, A.H., (1992), *Decline of Donnish Dominion: The British Academic Professions in the Twentieth Century*, Oxford: Oxford University Press.
- Harvey, D., (2005), *A Brief History of Neoliberalism*, Oxford: Oxford University Press.
- Hazelkorn, E., (2011), *Rankings and the Reshaping of Higher Education*, Basingstoke: Palgrave.
- Henkel, M., (2000), *Academic Identities and Policy Change in Higher Education*, London: Jessica Kingsley.
- Hicks, D., (2009), 'Evolving regimes of multi-university research evaluation', *Higher Education*, 57: 393–404.
- Hirsch, J.E., (2005), 'An index to quantify an individual's scientific research output' *Proceedings of the National Academy of Sciences*, 102 (46): 16569–16572.
- Holmwood, J., (2010), 'Sociology's misfortune: Disciplines, interdisciplinarity and the impact of audit culture', *British Journal of Sociology*, 61 (4): 639–658.
- Holmwood, J., (2011), 'TRACked and FECKed: How audits undermine the arts, humanities and social sciences', *Exquisite Life: Research Blogs*, available at: http://exquisitelife.researchresearch.com/exquisite_life/2011/03/tracked-and-fecked-how-audits-undermine-the-arts-humanities-and-social-sciences.html.
- Hovland, J., (2011), 'Numbers: Their relation to power and organization', in Saetnan, A.R., Lomell, H. and Hammer, S. (eds), *The Mutual Construction of Statistics and Society*, London: Routledge.
- Howe, G., (2005), 'Universities in the UK: Drowning by numbers', *Critical Quarterly*, 47: 1–10.
- Jacobson, H., (1983), *Coming from Behind*, London: Black Swan.
- Kelly, A. and Burrows, R., (2012), 'Measuring the value of sociology? Some notes on performative metricization in the contemporary academy', in Adkins, L. and Lury, C. (eds), *Measure and Value*, Sociological Review Monographs, Oxford: Wiley-Blackwell: 131–152.
- Kogan, M. and Hanney, S., (2000), *Reforming Higher Education*, London: Jessica Kingsley.
- Larsson, S., (2009), 'An emerging economy of publications and citations', *Nordisk Pedagogik*, 29: 29–52.
- Law, J. and Urry, J., (2004), 'Enacting the social', *Economy and Society*, 33 (3): 390–410.
- Lindgren, L., (2011), 'If Robert Merton said it, it must be true: A citation analysis in the field of performance measurement', *Evaluation*, 17 (1): 7–19.
- Lock, G. and Martins, H., (2011), 'Quantified control and the mass production of “psychotic citizens”', *Espaces Temps.net*, <http://espacestemp.net/document8555.html>.
- Lodge, D., (1984), *Small World*, London: Penguin.
- Lodge, D., (1988), *Nice Work*, London: Penguin.

- Lynch, K., (2010), 'Carelessness: A hidden doxa of higher education', *Arts and Humanities in Higher Education*, 9 (1): 54–67.
- Merton, R.K., (2000), 'On the garfield input to the sociology of science: A retrospective collage', in Cronin, B. and Atkins, H.B. (eds), *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield* (ASIS Monograph Series): 435–448, Medford, NJ: Information Today Inc.
- Neave, G., (1998), 'The evaluative state reconsidered', *European Journal of Education*, 33 (3): 265–284.
- Oakley, A., (1999), *Overheads*, London: HarperCollins.
- Parkin, F., (1985), *Krippendorfs Tribe*, London: Atheneum.
- Parkin, F., (1987), *The Mind and Body Shop*, London: Atheneum.
- Peck, J., (2008), 'Remaking laissez-faire', *Progress in Human Geography*, 32 (1): 3–43.
- Pelias, R., (2004), *A Methodology of the Heart*, Walnut Creek, CA: AltaMira Press.
- Power, M., (1997), *Audit Society*, Oxford: Oxford University Press.
- Power and Education*, (2010), Special issue on 'The Power of Journal Rankings', 3 (1). Available at: www.worlds/power/content/pdfs/3/issue3_1.asp (accessed 5 January 2012).
- Pringle, J., (2008), 'Trends in the use of ISI citation databases for evaluation', *Learned Publishing*, 21 (2): 85–91.
- Regini, M., (2011), *European Universities and the Challenge of the Market*, Cheltenham: Edward Elgar.
- Saetnan, A.R., Lomell, H. and Hammer, S. (eds), (2011), *The Mutual Construction of Statistics and Society*, London: Routledge.
- Shore, C., (2008), 'Audit culture and illiberal governance: Universities and the politics of accountability', *Anthropological Theory*, 8 (3): 278–298.
- Slaughter, S. and Leslie, L., (1997), *Academic Capitalism*, Baltimore, MD: Johns Hopkins University Press.
- Sparkes, A., (2007), 'Embodiment, academics, and the audit culture: A story seeking consideration', *Qualitative Research*, 7 (4): 521–550.
- Strathern, M. (ed.), (2000), *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy*, London: Routledge.
- Watson, D., (2010), *The Question of Morale: Managing Happiness and Unhappiness in University Life*, Buckingham: McGraw-Hill/Open University Press.
- Willmott, H., (1995), 'Managing the academics: Commodification and control in the development of university education in the UK', *Human Relations*, 48 (9): 993–1027.
- Winter, R., (1995), 'The University of Life plc: The "industrialization of higher education" ' in Smyth, J. (ed.), *The Changing Labour Process in Higher Education*, Buckingham: SRHE, Open University Press.
- Wright, S. and Rabo, A. (eds), (2010), 'Anthropologies of university reform', special issue of *Social Anthropology/Anthropologie Sociale*, 18 (11).