

# AUDIT CULTURE: WHY LAW JOURNALS ARE RANKED AND WHAT IMPACT THIS HAS ON THE DISCIPLINE OF LAW TODAY

---

KATHY BOWREY

## I INTRODUCTION

This article explores the shifts in research governance practices in Australian law schools that led to the production of ranking lists for law publications and reflects on the way these ranking lists were used from 2006 to 2012. As will become clear, this is a story of law gradually losing autonomy over design of the methodologies for research assessment, while seeming to continually and steadfastly oppose that movement. The example of journal ranking in law demonstrates how neo-liberal governance produces a myriad of sites of resistance, but how despite this, academic trust, goodwill and collegiality is readily co-opted into achieving management goals. In telling this story, I draw upon personal involvement in research policy work conducted by and for the Council of Australian Law Deans (CALD). I also examine public and private exchanges between legal academics and government personnel who oversaw implementation of research assessment policy, to explore the way academics continually sought to resist government interventions that would displace substantive scholarly assessments of legal research based on peer review. I explain why and how collaboration with journal ranking initiatives was nonetheless pursued, and reflect on some of the consequences of those decisions for legal academics today.

Part II begins with a brief overview of neo-liberalism and the audit culture in the higher education sector. Part III then documents the development of an audit culture in the discipline of law. I focus on the sector's demand that the discipline rank law journals in particular. Part IV presents a more personal reflection on the socio-cultural implications of what has passed, speculating on what

\* Faculty of Law, University of New South Wales.

Thanks to Margaret Thornton, and to Lesley Hitchens, Ben Golder, Gary Edmond, Mark Aronson, the referees, and Wendy Larcombe, for their helpful comments.

the longer term impact of the deployment of research assessment technologies in law may be.

The deployment of metrics to assess research quality in law is a subject about which many academics have strong feelings. Many academics are uncomfortable discussing their particular involvement in these processes, but this silence only further mystifies the governance practices. From experience, I think it is necessary to state at the outset that I am not arguing that legal researchers comprise a special group for whom there are no valid concerns about research quality. Nor is it my claim that there are no benefits to be had from highlighting questions such as, 'what does quality research look like?'. But, as Shore and Wright noted in relation to British audit culture of the 1990s,

The key question is not simply 'who is being made accountable to whom?' but rather, 'what are the socio-cultural and political implications of the technologies that are being used to hold people to account?'

There is not a good understanding of how auditing came into the discipline of law and this knowledge is essential to any evaluation of the implications of the management technologies being used today. In my view both participation in, and withdrawal from, engagement with governance technologies generate significant conflicts and strain collegiality. Accordingly, there is a significant problem in identifying how this form of governance can be effectively resisted with collegial and traditional academic values remaining intact.

## II NEO-LIBERALISM, AUDIT CULTURE AND THE ACADEMY

Neo-liberalism involves scientific business management based upon adoption of morally neutral criteria that inculcates a competitive, individualistic market mentality.<sup>2</sup> Audit culture is commonly referred to as a distinctive feature of neo-liberal governance. It is based upon the production of relevant metrics through which unit and individual performance is continually assessed and reported back to upper and middle management. Auditing leads to ongoing refinement of governance strategies, which may include financial rewards for high achievers and disciplinary consequences for poor performers.

In relation to higher education policy, neo-liberalism is commonly discussed in terms of the emergence of an audit culture and the

<sup>1</sup> Cris Shore and Susan Wright, 'Whose Accountability? Governmentality and the Auditing of Universities' (2004) 10(2) *Parallax* 100, 103.

<sup>2</sup> Wendy Brown, *Edgework. Critical Essays on Knowledge and Politics* (Princeton University Press, 2005) 40.

associated rise of managerialism within the academy. Margaret Thornton, for example, notes,

Managerialism is the transformative linchpin of the university that enables new knowledge to be mediated and harnessed by the state. Reflecting this crucial role, senior line managers (formerly 'administrators') have rapidly become the élite within universities, replacing professors. The task of line managers is to appraise academics regularly and ensure that they are 'productive', a process that needs to be demonstrated in performative terms.<sup>3</sup>

Governance practices that instil a neo-liberal market mentality erode traditional, scholarly values and power distributions. As managers become the new elite within universities, the authority and independence of members of the professoriate and other senior staff who were previously 'involved' in disciplinary based decision-making recedes, their voices becoming correspondingly muted and less influential. Discipline and sub-discipline specific knowledge that traditionally underpinned assessment of research quality is largely displaced by more generalised auditing formulas that are conducted centrally.

Quantitative data allows for comparison so that a comprehensive audit can be conducted, extending from the individual staff member, to a unit, group or department, to an entire school or faculty, which in turn feeds into a comparative ranking of each university and Australia world-wide. Public university budgets, which are widely accepted as inadequate, can then dedicate resources to support university research 'strengths' or concentrations of high-performers, through 'letting go' or re-assigning to 'teaching-only' posts the units, specialisations and staff that correspondingly appear to be 'unproductive' or 'uncompetitive'.

As a governance practice, audit culture is supposed to allow for greater accountability for performance at all levels of the institution. No one is exempt. However, as implemented within academia, it is inevitably imposed in a hierarchical manner on faculties, schools, departments and disciplines, by centrally administered budget allocations and decisions over staffing. In this context, it becomes exceedingly difficult to hold managers to account for their choices and the ensuing impact that decisions have at a disciplinary level. Critiques can be voiced but they have little impact on outcomes. In part, this is because a critique of the methodologies being deployed invariably speaks in terms profoundly different from those of the audit:

Both audit and positivism are rooted in the same assumptions about research and practice. Both exemplify the central characteristics of

<sup>3</sup> Margaret Thornton, 'Gothic Horror in the Legal Academy' (2005) 14 *Social and Legal Studies* 267, 271.

what Habermas (1972) termed instrumental or technical rationality ... representing a preoccupation with means in preference to ends, more 'concerned with method and efficiency rather than with *purposes*'<sup>4</sup>

Research assessment outcomes can be questioned or rejected as being seriously flawed.<sup>5</sup> However, such criticism often fails to affect decision making that follows because 'No policy maker wants to hear that things are messy, that the solutions are messy and partial at best, or that solutions to problems are uncertain'.<sup>6</sup> The managerial mandate is to make a strategic decision based upon the data, not to question the normativity of assessment mechanisms. The result is, as Shore and Wright recall, that university accountability is a one-way street. What constitutes poor management performance is seldom defined either in practice or in law.<sup>7</sup>

Neo-liberalism and audit culture is associated with a Foucauldian discipline of the self, with metrics engendering anxiety and insecurities so that academics continually scrutinise their own behaviour, choices and performance in terms of the norms and expectations of management. Shore and Wright argue that the logic deployed is one of hyper-surveillance: 'The rationality of the audit thus appears similar to that of the panopticon: it orders the whole system while ranking everyone within it. Every individual is made acutely aware of their conduct and performance is under constant scrutiny.'<sup>8</sup> Likewise, Thornton argues that 'A web of subinfeudation ensures that every person is answerable to someone above while overseeing someone below. In this way, governmentality is entrenched and normalised.'<sup>9</sup>

The fact that it is exceedingly hard to contest the facticity of research assessment is itself part of the *raison d'être* of a neo-liberal governance strategy. Anxiety helps instil compliance. Many sociological accounts discuss the anxiety-producing effects of audit culture within academia. Sparkes, for example, notes the psychic discomfort that comes with being a Director of Research and participating in these managerial exercises in the United Kingdom, while also trying to voice the usual disclaimers common in the humanities about the defects of the methodologies used in

<sup>4</sup> Phil Hodkinson, 'Scientific Research, Educational Policy, and Educational Practice in the United Kingdom: The Impact of the Audit Culture on Further Education' (2008) 8(3) *Critical Methodologies* 302, 309 (citation omitted).

<sup>5</sup> See for example, the analysis of the decision making that led to the closure of the Department of Sociology and Cultural Studies at Birmingham University following a poor result in the 2001 RAE in Cris Shore, 'Audit culture and illiberal governance' (2008) 8 *Anthropological Theory* 278

<sup>6</sup> Hodkinson, above n 4, 310.

<sup>7</sup> Cris Shore and Susan Wright, 'Coercive Accountability: the rise of audit culture in higher education' in Marilyn Strathern (ed), *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy* (Routledge, 2000) 110.

<sup>8</sup> *Ibid* 77.

<sup>9</sup> Margaret Thornton, *Privatising the Public University: The Case of Law* (Routledge, 2012) 209.

ranking exercises.<sup>10</sup> If it is believed that the academic performance of individuals and departments cannot be adequately judged with reference to numbers of publications in tiers of quality levels and other numerical impact factors, having to participate in such assessments, even if only to seek to defend colleagues, creates significant distress and anger, and reinforces a feeling of powerlessness.

Scholarly accounts of neo-liberalism and managerialism in the academy commonly focus on the higher level policy decisions of government, and the grassroots impacts on individual researchers. The description of neo-liberal governance often becomes quite abstract or site specific. However, neo-liberalism is more than a theory of management and control. It is constituted by practices that affect behaviour, conduct and identity at all levels of academia. Accordingly, it is important 'to move beyond paradigms that speak of neo-liberalism as a thing that acts in the world and focus instead on concrete projects that account for specific people, institutions and places.'<sup>11</sup> I would add to this a need to understand the attempts at resistance and push-back that occur when policies are sought to be implemented.

In the following section, I overview the policy demands made on the discipline of law to assist in refining research quality assessment measures and, in particular, the demand that law ranks its journals. As a participant in formulating part of the disciplinary response on behalf of the Council of Australian Law Deans (CALD) from 2006 to 2011, I cannot pretend to offer an impersonal or dispassionate account. Further, my participation in these processes came about largely as a consequence of my being available, through my changed employment circumstances. I was sufficiently concerned about the unfolding policies to not want to ignore what was happening, and felt unable, in good conscience, to excuse myself. However, in line with the theoretical literature described above, I also found it an intensely anxiety-producing experience throughout. That feeling remains in my attempt to try to account, in this article, for what transpired.

### III AUSTRALIAN RESEARCH ASSESSMENT EXERCISES

In my experience as an Associate Dean at UTS (2006–07) and at UNSW (2008–11) and through my involvement leading CALD responses to Australian Research Council (ARC) research assessment policies from 2006 to 2012, it was very rarely the case that law personnel embraced any sector management initiatives. Nor did they think criteria being used to assess research quality were

<sup>10</sup> Andrew C Sparkes, 'Embodiment, academics, and the audit culture: a story seeking consideration' (2007) 7(4) *Qualitative Research* 521, 527.

<sup>11</sup> Catherine Kingfisher and Jeff Maskovsky, 'Introduction: the Limits of Neoliberalism' (2008) 28(2) *Critique of Anthropology* 115, 119.

very sound or that assumptions being made were readily justifiable. For a long time law operated on wishful thinking, hoping the need to engage with sector policies in this area would simply all go away. Little effort went into thinking about developing appropriate performance measures or methodologies that might be relevant to assessment of legal research. Inevitably, those involved, including myself, saw themselves in a defensive role, seeking to protect the discipline and legal researchers as best they could against the ill-informed, often irrational methodologies of outsiders that could have serious consequences for (potentially) us all.

Journal ranking was always an initiative nested alongside many other policy measures, ostensibly designed to assess and improve Australian research quality, in a context of ongoing policy refinement overseen initially by the Hon. Julie Bishop under the Howard government and then by the Hon. Kim Carr under the Rudd and Gillard governments. Law journal ranking needs to be understood not as one decision related to one policy objective but as a metric developed over time under different regimes. Assessment that relies on journal ranking today is, arguably, only loosely related to the original policies that led to the genesis of the list.

### *A The Howard Government Research Quality Initiative 2003–07*

The Research Quality Initiative (RQF) was inspired by similar research initiatives such as the RAE in the United Kingdom and Performance Based Research Funding in New Zealand. Law became involved in 2006 when, on 15 November 2006, Minister for Education, Science and Training (DEST) Julie Bishop announced that the RQF would go ahead and released *The Recommended RQF*. The scheme was designed to assess both research quality and impact, with metrics playing a significant role in assessments.

From its early days, there had been many informal and formal discussions and trial studies about potential criteria. In 2006, a RQF Metrics Working Group recognised that in general ‘Metrics should be used to inform the peer review process rather than replace it’.<sup>12</sup> It was acknowledged in meetings conducted by DEST metrics experts that ‘best practice’ with metrics required a consideration of data reliability as determined through rigorous testing: testing for transparency, cost efficiency, positive behaviour impact (driving quality not mere productivity), the acceptability of the criteria to the community being assessed, and simplicity. A recognised problem facing the use of metrics experts was described in this way:

<sup>12</sup> RQF Quality Metrics Working Group, *Research Quality Framework: Assessing the Quality and Impact of Research in Australia*, (DEST, September 2006) 2.

Quite often I am confronted with the situation that responsible science administrators in national governments and in institutions request the application of bibliometric indicators that are not advanced enough. They ... want to have it 'fast', in 'main lines', and not 'too expensive'... the fault of these leading scientists and administrators is asking too much and offering too little.<sup>13</sup>

RQF bibliometric tools were not originally designed with the law or humanities in mind, thus any deployment of metrics in those disciplines required participation by relevant discipline representatives. Discipline-specific RQF Panels were implemented. These panels were expected to bear primary responsibility for assessment using a 'basket of quality measures', including metrics. Professor Hilary Charlesworth was appointed Chair of *Cluster II: Law, education and professional practices*. In addition to law, this grouping included criminology, librarianship, education, media and communications and social work.

It was anticipated from the outset that there would be push-back from some disciplines. It is hard to tell if law was originally thought of in this way, or whether experts in bibliometrics were simply less interested in law and other 'professional' disciplines, especially given their comparative size against the rest of the humanities. The Council of Australian Law Deans (CALD) was not consulted by DEST as a relevant peak body until July 2007, at a point when RQF methodologies for assessing research quality and research impact were already well advanced.

Details of the proposed scheme can be discerned from notes of the concerns that were raised at an initial consultation with law representatives, in a meeting held in Canberra and attended by available Deans, Associate Deans (Research) or nominees.<sup>14</sup> In general terms, law rejected the use of metrics such as ranking and citation to assess research quality. While the capacity to recognise the significance of policy-oriented legal research under an assessment of research impact was welcomed, concerns were raised about assessing impact with reference to 'adoption' of research by end-users. Adoption of legal research was not considered a sound indicator of the quality of the research, but more related to political fit.

Citation data was rejected on both practical and substantive grounds. Ranking of outlets was rejected on the grounds that any such proxy for quality would be unable to appreciate the complexity

<sup>13</sup> Van Raan, quoted in Colin Steele et al, 'The Publishing Imperative: The Pervasive Influence of Publication Metrics' (2006) 19(4) *Learned Publishing* 277, 279.

<sup>14</sup> Following this event CALD appointed an RQF Working Party. It was comprised of myself and Professor Mark Israel (Flinders University). Formally the work included writing briefing notes on sector development for CALD, drafting responses to DEST requests for information and drafting numerous submissions to DEST as required.

of factors affecting the choice of publisher and journal in law, including the specificities of jurisdiction, the need to relate to professional and interdisciplinary audiences, differences between 'core' areas and specialisations and respective size of different areas of legal research. The unfairness of retrospectivity was also raised as a concern. More generally, there were concerns over the proposed size content of 'evidence portfolios' required, which included criteria for selection of 'best four' publications. Questions were raised about the respective 'weighting' of books against journal publications and what considerations addressed article length, which might range from 5,000–6,000 words at the low end to 8,000–12,000 for a 'standard' journal article or chapter, with up to 20,000 words for some works at the extreme end. Would these all be counted as 'equal' works in metrics algorithms? Commentators were indifferent to the inclusion of refereed conference papers. Given that the indicators of quality could affect institutional 'research active' assessments and future researcher behaviour, there was discussion of Given that the indicators of quality could affect institutional 'research active' assessments and future researcher behaviour, there was discussion of the status afforded to original law reform and policy work. Law reform and policy work did not amount to reportable publication under current sector rules, but it could count for the purposes of measuring 'impact' under the ERA. The presumption was that law reform work was always derivative or minor, based upon pre-existing original scholarship in a journal or book publication. However in reality some law reform and policy contributions were original works produced for the organisation that invited the contribution. Where so, this effort needed to count as a research publication, as well as evidence research impact. It also became apparent that there would be overlaps within Cluster 11 (such as between criminology and law) and between the Professional Cluster and the humanities more generally. Was multidisciplinary work to be discounted when it came to assessment in two cluster areas, or double counted? The appointment and role of international advisors to 'mediate results' and 'overcome knowledge gaps' was also discussed, mainly leading to advocacy for particular names. More generally, there was a concern for the time-frames envisaged, the scale of the undertaking sought and for the small commitment of resources and time allocated to the entire process. Overall, there was much scepticism and little support for assessment of law by metrics at all.<sup>15</sup>

<sup>15</sup> This view was soon supported by a report of the British Academy, which noted while peer review did pose potential problems with bias and subjectivity, it remained the best available measure for assessing quality and that metrics needed to be resisted. The British Academy, *Peer Review: The Challenges for the Humanities and Social Sciences* (2007) 29.

Nonetheless, in August 2007 CALD was asked by DEST to begin ranking law journals and publishers. A CALD document circulated the candid views of 30 law schools. It revealed that ranking was only ever entertained by Law Deans as a last resort option, a task only to be entertained because it was believed that law had been given 'little choice'. Deans were concerned for what non-compliance with the DEST request could eventually mean for sector funding of legal research, also about the potential consequences of law being perceived as obstructionist within home institutions. The formal advice from DEST to the discipline was that law would be the only discipline not ranking. In hindsight this proved a little misleading. Other disciplines also had difficulties with the notion of ranking publication outlets. Law was further advised that, without the indicator of ranking, RQF Panel Assessors would have nothing to go on but peer review. This was in a context where it had already been privately acknowledged that reading more than 25 per cent of the submission was unlikely to be possible. A significant dilemma now existed. Was the potential arbitrariness of assessment based on a cursory reading of some work by a handful of assessors going to be improved by including an across-the-board application of ranking metrics? DEST further advised that the lack of metrics was likely to have detrimental consequences for law, although there were no details of how metrics factored into anticipated institutional funding allocations in a way that was reflective of quality assessments.

At this point, the CALD strategy settled on trying to retain as much autonomy as possible, at discipline level, over all the assessment processes and in particular over the makeup of potential assessor and advisor lists. An interest was also taken in providing the ARC with more accurate and necessary information about difficulties in assessing legal research, especially given the diversity of legal research outputs and audiences. In the context of assessing research impact there were concerns about who should be considered an 'end-user' of legal research and about the limitations of referring to citation in judgments, reports or committees alone to demonstrate impact. It was argued that in appointing advisers and assessors, the ARC needed to be aware that law saw itself as a 'professional discipline' in a very qualified sense; that members of the profession were not necessarily relevant assessors of research quality or impact, and a fair assessment process would need to be sensitive to different approaches and styles of legal research and include a balance of expertise sensitive to differences between doctrinal, policy-oriented and theoretical legal research.

The difficulties of comparing different kinds of legal research soon struck home at institutional level as the RQF required nomination of 'best four' outputs of individual researchers. What did *best* mean? Potential criteria included a consideration of perceived inherent

qualities of the work, such as whether it had impressive depth in coverage of subject matter, originality, currency and the work's scholarly reception. Given that short cuts may be taken to assess quality, more strategic indicators were also considered. These included outlet reputation, author reputation, citation, likely familiarity with the work, whether the work was controversial, likeability to potential reviewers and so on. It was suggested by experts in bibliometrics that the difficulties in determining *best* criteria pointed to broader problems with the subjectivity that is inherent in peer review, and thus the need for law to embrace metrics and rank publication outlets was underscored. This view however, was strongly countered by the results from RAE law assessments. For example, the 2001 RAE Law Panel's Overview Report stated that:

Work of internationally recognised excellence was found in a wide range of types of outputs and places, and in both sole and jointly authored works (the Panel adhered to its published criteria in allocating credit for joint pieces). First-rate articles were found in both well-known journals and relatively little-known ones. Conversely, not all the submitted pieces that had been published in 'prestigious' journals were judged to be of international excellence. These two points reinforced the Panel's view that it would not be safe to determine the quality of research outputs on the basis of the place in which they have been published or whether the journal was 'refereed'.<sup>16</sup>

In 2005–06, British legal scholars had successfully resisted attempts to rank journals, pointing to the size of the discipline and the high degree of specialisation within it making objective evaluation of law outlets extremely difficult. They pointed to past RAE assessments that showed the publication outlet was an unreliable indicator of quality. However, there was also ongoing disquiet from some quarters in the UK about the resourcing and reliability of peer review. Australian bibliometrics experts sought to impress their international peers by demonstrating how metrics could be incorporated to improve efficiency and reliability of UK assessments, including in law and the humanities. The outsourcing of Australian research metric technologies to New Zealand was also discussed as a possibility.<sup>17</sup>

### B *The Rudd–Gillard Government Excellence in Research for Australia Initiative (ERA)*

On 21 December 2007 the Rudd government announced the abandonment of the RQF and the development of the Excellence

<sup>16</sup> RAE 2001 Overview Reports from the Panels', Unit 36, Law (2001) <<http://www.rae.ac.uk/2001/overview/>>.

<sup>17</sup> This observation is based on conversations and readers' postings on blogs by relevant Australian personnel at the time.

in Research for Australia (ERA) initiative, led by the Australian Research Council (ARC). Under the ERA, law was assigned to a Humanities and Creative Arts (HCA) Panel. This scheme did not include an impact measure but retained reliance on peer review of 20 per cent of outputs, with metrics that were to be applied to the whole submission of reportable publications. The latter process required the ranking of law journals across four bands: A\* (top 5%), A (next 15%), B (next 30%), C (next 50%). Rankings were required to be internationally valid in order to place published Australian research within a comparative international context.

There were major logistical challenges that frustrated the creation of a law list from the outset. This included there being no starting list of peer-reviewed law journals available. In June 2008 the ARC surprised and outraged many when it released preliminary rankings of law journals. These were later confirmed to be based on Washington and Lee Journal Rankings.

While there had been some consultation with other learned academies and peak bodies, neither CALD nor the Academy of Law had been formally consulted, nor kept abreast of developments. Relations with the ARC bibliometrics personnel were already strained, due to the earlier exchanges over law's participation in RQF journal ranking. With law unable to accept that ranking was appropriate for the discipline, and the bibliometrics personnel unable to accept that it was not, there was little common ground for further engagement over the development of methodologies for ranking under the ERA.

Initial objections to the rankings made to the ARC by concerned legal researchers and some peak specialist bodies, such as the Australian and New Zealand Society of International Law (ANZSIL), included evidence that they did not meet the ARC's own criteria for validity. They were not international rankings as they did not include many, and major, journals from the United Kingdom, Europe and Commonwealth countries, and there were only two non-US law journals in the top 198. The ARC did not appear aware that US journals generally had different review processes from ours, and historically law had not accepted US student-run journals as 'peer-reviewed'. The attention of bibliometrics experts was drawn to additional analysis and literature that discussed the limitations of the list and to the fact that for rankings in law, there were differences in publication markets and scholarship practices between the US and other jurisdictions. It was also pointed out that publication in a US journal may not be at all relevant to Australian researchers, that there were problems with the Washington and Lee list accounting for research specialisation in law, that the role of general law journals differed from that of specialist journals, and so on.

From emails and phone calls I had with key personnel, I was left with the strong impression that release of the law list was a provocative step designed to force law *to accept the overall concept of ranking journals*, through being forced to engage in ‘modifying’ a list. If so, this objective was achieved: from this time on there was no longer any real discussion of rejecting rankings but rather only discussion of how to ‘improve’ a bad situation. I felt senior bibliometrics staff were indifferent to law’s concerns because we were a small and self-contained discipline and were taking up too much of their time, when they were subject to very tight time pressures within the department. Peer review was generally distrusted, at least in part because it was a process largely opaque to their own research expertise and because their bibliometric expertise was not really necessary to conduct it. In resource terms, they felt peer review was inefficient as it involved duplication of effort to subject an already peer-reviewed publication to further additional peer review. In bibliometric terms, assessing the overall journal reputation and preferably also citations was felt to be sufficient, and more objective. Our queries about how areas of such diverse specialisation in law could be compared, and the problem that any list would indirectly create a hierarchy of research importance, were ‘too difficult’ for them to deal with. In terms of more junior personnel, as their instructions were to implement metrics, the presumption was always that only minor tinkering through adding a few missing journals and some help in assigning correct field of research codes to multidisciplinary journals was really required. It was pointed out to me that there would be an opportunity for individuals to make the case for individual changes to inappropriate rankings and thus corrections would be made to the list as part of the process at a later stage, handled by them.

I was told that the Washington and Lee based list ‘stood up’ to scrutiny according to bibliometric criteria of validity. There was no attempt made at all to justify the methodology as appropriate for Australian legal research or for the purpose it was to be put. The presumption was that as the ERA was to involve an international, comparative assessment, it was inappropriate to take local concerns about local publications too seriously, lest it be pandering to domestic insecurities and self-interest. The draft list was presented as a fait accompli based on the departmental belief that there was to be no further negotiation about the issue and that ranking of law journals would be implemented as part of the ERA with or without further involvement from law.

There were further opportunities to communicate with the ARC over the implementation of the ERA. Questionnaires were sent to each institution requesting feedback on particular assessment features. However there was little capacity to organise a well-coordinated response from law through this avenue. There was no current list of law Associate Deans or Directors of Research to

facilitate easy communication between law schools. The nature of law's objections to the draft list was not necessarily understood within home institutions, and central Research Offices were generally disinterested in pursuing law's 'special' case.

It was apparent that, without objections being forcefully conveyed at the most senior levels, there would be little possibility for anything more than an ad hoc revising of the list, to be primarily undertaken by the ARC or persons they co-opted to assist. There would be no transparency or accountability. Rumours began to circulate that some individuals at Group of Eight (Go8) institutions thought they could step in and take on the task of fixing the list for the ARC. The then Chair of CALD, Prof William Ford (UWA), discussed the problem with significant members of the judiciary and profession. Numerous communications were made to the ARC personnel and to then ARC Chief Executive Officer, Prof Margaret Shiel. This intervention resulted in the concession that CALD could, in consultation with ARC bibliometrics experts, undertake its own ranking exercise. It was a requirement that the CALD list be based on modifying and correcting the existing ARC list and that it involve further consultation with legal researchers and law specialist bodies.

The July CALD Meeting conveyed the view that law favoured genuine peer review, but supported metrics and ranked publication outlets 'so long as those metrics are appropriate to and supported by the discipline'.<sup>18</sup> CALD noted that the draft ARC ranking based on the Washington and Lee list failed to provide an adequate or relevant international benchmark. CALD proposed instead to develop a ranking in consultation with the ARC that would meet its deadlines. A CALD Steering Committee of Law Journal Ranking was then appointed, made up from an open invitation to Associate Deans (Research).<sup>19</sup>

#### *The CALD Journal Ranking List*

The list provided by the ARC had approximately 1,400 outlets but very few journals that Australians would seek to publish in, especially within the top tiers. The CALD ranking methodology triangulated information from a number of sources to substantially revise this list and produce a first-cut CALD revised draft. This process took into account:

<sup>18</sup> Council of Australian Law Deans Meeting 10 July 2008, 'Resolution #2', [2].

<sup>19</sup> The meeting was held at short notice. It was poorly attended. The Steering Committee comprised of volunteers included Kathy Bowrey (UNSW), Lesley Hitchens (UTS), Kit Barker (UQ) and Richard Johnstone (Griffith). Brad Sherman (UQ) was later co-opted to assist with some consultations. CALD provided funding for a part-time administrator to assist with data management, and UNSW provided additional technical research support.

- feedback from 22 Australian law schools on all journals on the list including identification of missing journals to be added;
- information provided by 82 journal editors in line with a template produced by the ARC about acceptance and rejection rates and other details editors thought relevant; and
- 34 general submissions received from law schools and interested individuals. While there was a rough consensus over rankings of approximately

85 per cent of the Australian journals, and information was received about many of the more prestigious US journals, scant information was received about a very large number of US general law journals that dominated the Phase One list. On advice from the ARC, it was not open to remove US journals from the list if these outlets appeared as 'equivalent to peer-reviewed' on the *Ulrich's* Periodical Directory, regardless of whether this was the case. Former RQF Panel 11 Chair Professor Hilary Charlesworth (ANU) assisted with revision of the ranking of these US general law journals. In the absence of any feedback, reference was made to the US News Ranking of Law Schools.

In many cases, feedback was most enthusiastic and well organised by persons who had a vested interest in particular journal rankings. It was not possible to manage real or potential conflicts of interest that arose from the consultations conducted in Australia (except those concerning members of the CALD Steering Committee). ARC advice was that 'errors' that would become apparent over time could be corrected through incorporation of feedback in Stage 2, which would be provided by specialist bodies and independent international reviewers. Twenty-five specialist and professional bodies were approached and asked to participate in reviewing the second cut list. In addition, there was liaison with the Australian and New Zealand Society of Criminology (ANZSOC), which led to a redistribution of some journals to criminology and others to law and an assignment of both codes for journals that were considered interdisciplinary. All this occurred within tight time frames, limiting opportunities for extended discussion and engagement.

A list of 100 potential eminent international reviewers for particular specialisations was devised and circulated to 22 senior Australian academics for comment as to appropriateness and for additional suggestions. All those approached were highly regarded by all Australian peers. They were also invariably people who had experience in research assessment and a reasonable knowledge of Australian or Commonwealth legal research culture and journals. The 61 people ultimately approached, two to three for each area of specialisation, included representatives from the United Kingdom, Ireland, Singapore, Canada, New Zealand and the United States.

The responses to international invitations to participate in the CALD/ARC process are worth considering in some detail. Despite being sympathetic to the pressure to have ‘international endorsement’ of draft Australian rankings, the two peak UK bodies, the Society of Legal Scholars (UK) and UK equivalent to CALD, the Committee of the Heads of University Law Schools, declined to participate on the basis that journal ranking is a flawed measure of research quality. As President of the Society of Legal Scholars, Professor Fiona Cownie, wrote

You may be aware that there is a widespread view among academic lawyers in the UK that a journal ranking table is not appropriate for the academic discipline of Law. This is a view with which the Society of Legal Scholars concurs. Our view is reinforced by the approach of every Law panel there has been, in successive U.K. Research Assessment Exercises. A consistent finding of these panels has been that very high quality work is to be found across a very wide range of journals, and not just those that informally have been seen as the most prestigious, and that not all the work in such journals is of very high quality ....

There are particular variables affecting legal scholarship not necessarily replicated for other disciplines. Writing may be jurisdictional, comparative or international; may be directed to practitioner or academic audiences or both; may be doctrinal or socio-legal; and may be highly specialised. As a result, there is a very wide range of legal journals; ranking would be likely to devalue certain kinds of high quality legal publication, such as those related to the law of a small jurisdiction or in particular specialist areas. This would be unfair to the scholars and hinder the proper development of the law in those jurisdictions or areas ....

Given the view of the Society I do not feel it would be appropriate for me to participate in the Australian exercise.<sup>20</sup>

There was a similar response from the Committee of the Heads of University Law Schools. Ranking journals was not considered as a ‘credible’ tool for assessing research quality in law.

A significant number of distinguished academics also declined to participate and either asked for their comments to be passed on to the ARC, or only agreed to participate on condition of their objections being passed on. These included a range of comments compiled below<sup>21</sup>

... I believe the exercise to be fundamentally flawed ... The reason is that there is no recognised hierarchy of journals, so that very good articles are to be found in journals that are not particularly well known and, sometimes, relatively weak ones get into quite ‘prestigious’ journals. I think I can say that the experience of two RAEs only served to confirm the correctness of the UK Law Panel's decision to pay no attention whatever

<sup>20</sup> Letter from Fiona Cownie, President, Society of Legal Scholars to CALD Journal Ranking Steering Committee, Nov 2008.

<sup>21</sup> Feedback received by CALD Journal Ranking Steering Committee, September, 2008. Identification of authors has been removed in order to preserve the confidentiality of individual participants.

to where a piece was published but to concentrate on whether or not the individual piece was of good quality...More fundamentally, all that an exercise ... I think you are seeking to carry out can do is to give you a historical record of which journals carried the best articles during the last x years. What certainty is there that the best pieces will appear in the same journals in the next period? In the absence of an established ranking scheme, none whatever.

I regret that this is something I would not like to do. I am strongly opposed to these government inspired exercises which in the long term will be only detrimental to academic freedom.

I think this whole process is misguided and a waste of time. While one can probably assign relative value to individual articles (though even then, it would depend on what counts as value, and the ranking would be heavily influenced by the biases of the person doing the evaluation), to try to do so for entire journals would inevitably undervalue some articles and overvalue others. I could not in good conscience participate in such a process

The more I looked at this, however, the more misconceived I felt the whole exercise to be. The exercise of attempting ratings convinced me it was like comparing apples and pears. Nevertheless, I shall comment on a few journals for what it is worth, but first I must make some general comments

- [Ranking methods] favour the general over the specialist, yet monitoring of specialist journals is often much more thorough than in general journals.
- It is easy enough to pack an editorial board with field leaders without any of them actually contributing to the work of the journal.
- Ranking journals assumes that ranking will be static. Ranking may indeed become self-fulfilling, in that researchers will submit to high ranking journals, but editorial teams will change and quality changes too
- Ranking journals gives much too much power to a few self-perpetuating groups of editors.
- Ranking journals discourages innovation, inter-disciplinarity and development.

In June 2009, the ARC published a revised list (the HCA list) which was primarily based upon all feedback received by CALD. The HCA list was utilised for an ERA Trial conducted in the second half of 2009. Unexpectedly, in September 2009, the ARC conducted another round of feedback on journal ranking in anticipation of ERA 2010. It appears a number of organisations were invited to comment. Journals unhappy with their own ranking, or that of a rival journal that was ranked higher than theirs in the HCA list, were also given an opportunity to participate in providing additional feedback. This information was incorporated into a new revision by the ARC, in most cases by complaints being accommodated. There was no involvement of CALD or former members of the Steering Committee

in this stage of the process. A final stage consultation was then undertaken in late 2009. A revised list was confidentially circulated to a number of invited individuals for comment and final review. The ARC was happy for the original CALD rankings to be reinstated as these rankings were supported by a range of data sources. However, it was not possible to identify all the changes made to such a large list in the time frame given and from the information provided.

The final ERA 2010 list was published in February 2010 and utilised for ERA 2010. It contains significant differences from the earlier HCA list. In the final version, a significant number of Australian law journals were removed from the list, possibly because of doubts about their peer review status, however comparative US journals remained unaffected. A large number of US general law journals and the *Griffith Law Review* were upgraded. Some specialist law journals, of interest to Australian legal researchers and with strong support for their existing ranking, were downgraded. These unanticipated changes caused some consternation. Protest to the ARC over the methodology utilised in Phase Three and Four consultations by legal researchers echoed broader concerns from other disciplines over the longer-term consequences of the ranking for Australian researchers.

These concerns centred on the way rankings would deleteriously affect the health of Australian publications in the humanities, where financial viability was often already precarious. Some publishers, anxious that a journal was already financially marginal, sought to put the titles on the market. Concerns also arose because the proportion of publications reported in A and A\* journal lists was influential in institutional ratings for ERA 2010, and this fed into institutional assessments of research strengths. Some institutions began evaluating research areas and individual research performance with reference to quantum of publications in various bands.

Based on these considerations, in May 2011 in a Ministerial Statement, the Hon. Kim Carr announced the abandonment of journal rankings. He noted, 'There is clear and consistent evidence that the rankings were being deployed inappropriately in some quarters within the sector, in ways that could produce harmful outcomes and based on a poor understanding of the actual role of rankings ... One common example was the setting of targets for publication in A and A\* journals by institutional research managers.'<sup>22</sup> It was however, clear from the outset that this would be a consequence of the policies set in train, and well understood at departmental level. The back-peddalling by the Minister shows the effectiveness of loud voices from the sector, but for reasons explored below, it did not entail much of a victory.

<sup>22</sup> Jill Rowbotham, 'End of an ERA: journal rankings dropped' *The Australian* 30 May 2011.

In ERA 2012, ranking ERA assessors were provided with data about the most frequently published in journals, without reference to past journal rankings.

## 2 *Law Research Assessment Transformed*

Throughout this period, the discipline remained publicly opposed to journal ranking and generally deeply suspicious of any shift from traditional discipline-centred modes of assessment. Nonetheless, for instrumental reasons, law continually conceded ground. There was a strident defence of what was understood as traditional academic values, particularly centring on a defence of peer review as the primary tool of assessment. However, the claim was reduced to that of peer review remaining as a part of the equation used within a centralised assessment process. With short timelines, heavy workloads and inadequate classification and knowledge about relevant expertise, the character of peer review was fundamentally compromised and changed in the process. What is involved, in reality, is not ‘traditional’ discipline-centred peer review, but something else. The practice of peer review has been transformed by a technical rationality.

Hodkinson argues that the accountability strategy of neo-liberalism is one that seeks ‘to undermine the independent power of professional bodies, which were seen as promoting their own self-interests.’<sup>23</sup> In effect, what has occurred in law is an undermining of the authority of the discipline. This shift places academics in an invidious position. One can adopt a managerial role and take on the role of policing the discipline through participation in these processes. This can bring real career benefits, given the institutional esteem that flows from being able to show one is engaged in higher level management of one’s discipline. Or, one can eschew that role out of a commitment to more traditional knowledge-based status and collegiality.<sup>24</sup> Either way, the culture of the discipline is transformed as discipline-centred knowledge and expertise loses influence and impact over assessment of research quality in law.

## IV WHERE TO NOW?

### A *Ongoing Impacts*

As part of yearly compliance with the *Higher Education Research Data Collection* (HERDC) that is used as part of university funding formulas, institutions generate significant reference data annually about the publications of their employees. The ARC law journal ranking list and revised reiterations of it made by interested parties is now being used, or sought to be used, as a proxy to assist measurement

<sup>23</sup> Hodkinson, above n 4, 306.

<sup>24</sup> This dynamic is described by Shore and Wright, above, n 1, 75.

of *individual research performance* based on reported publication *outcomes*. The current industrial uses of ranking data have little to do with the original aims of improving research quality or of allowing for comparative assessment of institutional performance. What began as a quality assessment exercise that law participated in at the instigation of the ARC and due to pressure from home institutions has thus morphed into a management tool that plays a direct role in industrial processes affecting employment terms and conditions. That this next stage would unfold was entirely predictable, given the governance logic of neo-liberalism as described in Part II. Labour cannot be transformed in line with managerial goals, as the data indicates is required in the sector, unless knowledge workers are impacted at the coalface by being rewarded or punished as the data suggests is merited.

In 2012, concerned to document what had occurred and to inform the various institutional uses being made of the journal ranking data, CALD commissioned a report to better explain specifics of the older ranking list.<sup>25</sup> The report includes significant data analysis of the journals ranked, including information about known problems with the rankings of general and specialist law journals. In reading the information on law journals contained in that report, it needs to be noted that the substantive data underlying most ranking lists now in use is five years out of date. The data primarily reflects perceptions and other details of publications dated before 2008. Though revised data sets may be discreetly updated or doctored, any journal ranking lists now being used to assess individual research performance in law are unlikely to be defensible based on *any* accepted bibliometric methodology. An arbitrary ranking thus creates ‘presumptions’ of relative researcher merit with validity of the data simply presumed because the same data set is consulted to assess all.

The ability to ‘run the data’ thus sets in train presumptions of relative merit that then need to be contested and rebuffed by individuals, affected groups and their advocates as best they can, based on ‘supplementary’ information to explain ‘poor’ publication choices, a ‘different’ career trajectory and ‘career disruption’. The centrality of data collection and reliance on metrics within university processes means that so long as ranking data is available to assess legal researchers, it will continue to erode and undermine reliance on more traditional indicators of research performance in law — peer review, based on a substantive and well-informed assessment of a body of work through referee reports, book reviews and other established indicators of reputation and standing. Where it can be made to count at all, discipline-specific assessment of candidate

<sup>25</sup> Chapters 3 and 4 specifically address journals. See Kathy Bowrey, *Assessing Research Performance in the Discipline of Law, 2006-2011*, (2012) <[http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=206080](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=206080)>

performance inevitably comes to be reduced to just another ‘variable’ to be weighed against the centrally produced data, having a taint of ‘subjectivity’ hanging over it.

What is being normalised and rewarded here is the career path and profile that most neatly fits with the normativity embedded in the metrics — repetition of regular output drawing on an already familiar knowledge base, published in ‘safe’ journals, in not-too-long articles, produced with the benefit of minimal ‘disruption’ by teaching, administration, professional and other collegial contributions, political engagement, health or family life. While research assessments purport to reward quality, there is no attempt to account for the intellectual and labour conditions under which quality work may be produced. This would require a more detailed consideration of the whole individual with reference to their career story, longer-term goals and ambitions, and attention to the particular conditions, challenges and values of the research field(s) they work in. However, as noted above, audit culture embodies a technocratic rationality and is not concerned with the ends or purposes achieved.

The research metrics being used in law will inevitably deliver arbitrary results in assessing quality. However, a technocratic rationality seeks to suppress that reality and naturalise the arbitrariness of the new order by requiring researchers to fall into line with the current rules of the assessment game.<sup>26</sup> In the face of these kinds of dynamics, dispirited individuals can seek to withdraw, as far as that is possible. It may be possible to raise objections when the opportunity arises, but in doing so, the individual risks being seen as *difficult naïve* or engaged in *special pleading*. Raising problems can embarrass and humiliate senior colleagues who can’t defend the processes and generally don’t seek to, but whose own performance and relationship with management is judged on adherence to the management norms. Anxiety is not just felt at the bottom but permeates throughout all levels. For those who retain an allegiance to older scholarly values, it produces internal identity conflicts. These are impossible to resolve, with hard choices requiring the sacrifice of scholarly judgment, as friendships and careers are continually being placed in front of you. As non-compliant managers are moved on by their superiors, it is strategic to settle for the occasional compromise or success in blocking a highly unpopular decision and through that avenue, help restore the confidence of colleagues. It is naïve to hope that a benign or well-meaning senior manager can mediate every likely injustice. One is more likely to go into bat where there are existing personal ties, strong networks and management imperatives that make a person inherently of more value to the manager or institution. While law schools have always been hierarchical and favouritism is hardly

<sup>26</sup> As Bourdieu notes, ‘every established order tends to produce ... the naturalization of its own arbitrariness’. Shore, above n 5, 279.

a new factor in facilitating career success, current practices do little to improve things. Arguably, the new modes of governance entrench these problems, making them increasingly intractable. Over time, and at all levels of management, we are pressured into being more *selectively* collegial and only taking on roles, fighting for people or issues if they are *worth the hassle*, when *the time is right*, when people are *more likely to be receptive*, when a favour is needed in return or when there will be little repercussion. Inevitably, as a matter of mental health and personal wellbeing, one is forced to be more and more selective in choosing what to resist.

### B *Being More Strategic*

The technical tools deployed and now governing the sector are presumed to channel academic energies into serving the needs of a deregulated economy and the gods of productivity and efficiency. However, different disciplines have different roles to play in a neo-liberal economy. In this regard, the ‘one size fits all’ auditing of the sector, which treats all disciplinary knowledge bases as reducible to the same measure, becomes problematic.

Law is still often called a ‘professional’ discipline, but it is not the needs of the legal profession per se that legal academia serves. The professionalisation of legal education has led to a much more nuanced and sophisticated understanding of the relationships between law, the profession and the needs of society.

Law is an unusual discipline whose very character is affected by neo-liberalism because law is a tool of neo-liberalism. This makes us different from other disciplines because our discipline has to cope with significant direct pressures and tensions that stretch the possibility of what law is. In a neo-liberal world, law and legal processes become more and more diffuse, diverse and fragmented. The methodologies for knowing the law correspondingly shift, morph and differentially fail to relate to desired social and economic policies. One consequence of this is that as law becomes more diffuse, it is correspondingly harder and harder to categorise correctly and to assess retrospectively the associated legal scholarship. What may have been a sensible categorisation for an earlier time now appears to be more random, or an accident of history.

Metrics-based research assessment and evaluation of individual researcher merit is very, very heavily dependent upon the underlying research classification. Law cannot be assessed as ‘law’. It needs refinement into smaller units of assessment. However, government

<sup>27</sup> See Bowrey, ‘Part Two. Research Assessment Codes’, *Ibid.*

<sup>28</sup> That there were probably more law journals on the ARC law journal ranking list than there were actually legal academics employed in Australia is one reflection of this.

classification schema for legal research is notoriously poorly equipped and ill-conceived.<sup>27</sup> The discipline of law has always been extremely broadly based, diverse and diffuse.<sup>28</sup> With the changing character of neo-liberal law, it will be difficult for some time to develop new sub-disciplinary classifications, and the attempt to impose new ones is politically fraught.

This reality of the world we now live in will only further strain the ability of the technologies that are being used to effectively capture legal research. It is one thing to use arbitrary measures to assess research quality. But it is another to entrench or naturalise those tools when serious doubts exist about the capacity of those measures to deliver the desired policy goals, which in law is to serve the prospective needs of government and society. Highlighting this problem is one possible avenue for more productive, if unsettling, discussions about the deployment of research assessment tools in law today.

Without an acceptable mode of classification for legal research, not only can the appropriate assessors not be found, it is impossible to judge whether the larger government investment in research will bear fruit. Thus there may be some benefit in pursuing a strategy of turning neo-liberal discourse back on itself through constantly evaluating and interrogating the effectiveness of policies deployed to assess performance in law. Though reflexivity can be a double-edged sword, I think we can do a lot more as a discipline than we have in the past to combat the arbitrary processes to which we are now subjected. We need to begin by pointing out some of the unique facets of our discipline and the ways in which we are already accommodating the need to change by responding to the world we are in. Discussing what legal research is today, and revisiting what the various strands are, is something that could be worth doing, as well as taking the time to reflect on our activities as Australian legal researchers, our politics and the practices we engage in. It is one way we can begin to reassert more discipline-centred authority and expertise over the definition of legal research and its political value. It could form the basis for a more positive mode of resistance than simply complaining and complying, given that to date, this has been rather unproductive.