

limitations of particular measures and sources. He notes that Elsevier's SCOPUS service, from which HEFCE provides citation data to panels as 'additional information', does not cover the open access arXiv repository of 907,086 e-prints in physics, mathematics, computer science, quantitative biology, quantitative finance and statistics.<sup>71</sup> Nevertheless HEFCE's intention to press on regardless is plain. '*Expert review of the outputs*', it thunders as it rides into the valley of Death, '*will remain the primary means of assessing them*'.<sup>72</sup>

### 2.3 EVALUATING OUTPUTS: 'AN APPROPRIATE BREADTH OF RESEARCH EXPERTISE'?

If individual REF panellists' personal judgements play such a critical part in appraisals of outputs, then the least that can be asked is that – in the words of HEFCE's guidelines – 'The subpanel members and additional assessors should provide *sufficient breadth and depth of expertise* to undertake the assessment *across the subpanel's remit* (including as appropriate expertise in interdisciplinary research and expertise in the wider use or benefits of research)'.<sup>73</sup> Not only does REF 2014 lamentably fail to satisfy this requirement, there was no way in which it could ever have hoped to do so. This is not just a matter of this or that area of research not being adequately covered by this or that subpanel. The problems go much deeper. At issue is what constitutes peer review.

It is a curiosity, to begin with, that an exercise that purports to evaluate the 'originality, significance and rigour' of outputs 'with reference to international research quality standards' should leave their appraisal wholly in the hands of people based in British universities (and predominantly of older, male, white people based in British universities at that).<sup>74</sup> Unless we make the ludicrous (though not uncharacteristically older, male and white British) assumption that UK-based academics are top of the class in everything they do, this must mean that in at least some areas the REF is denying itself the opportunity of making use of the best international expertise available. The hubris involved here is all the more striking in that REF *defines* the quality of research *entirely in terms of its international reputation* – with 4\* research being 'world-leading', 3\* 'internationally excellent' and 2\* 'internationally recognised'.<sup>75</sup> Research that is only

<sup>71</sup> Coles, 'The apparatus of research assessment'.

<sup>72</sup> REF 2014: *Panel Criteria*, para 121. Emphasis added.

<sup>73</sup> REF 2014: *Units of Assessment*, para 55 (b). Emphasis added.

<sup>74</sup> See REF 2014: *Analysis of Panel Membership* for a detailed breakdown.

<sup>75</sup> REF 2014: *Panel Criteria*, para 37. Emphasis added.

'nationally recognised' gets the lowest grade of 1\*, yet HEFCE sees no inconsistency in recruiting the people who will make these judgements of relative international standing from a single national pool.

The Roberts report on the 2001 RAE recommended maintaining 'dependence upon expert peer review to identify the best research' as evaluated by 'panel members recruited from within the research community', but added the important rider: '*but not necessarily all UK-based academics*'.<sup>76</sup> Two articles published by way of an obituary for the RAE in *Times Higher Education* on 30 November 2007 shed further light on this perennial problem. 'International assessments are still being made by national assessors, first and foremost', complained Alex Danchev, then Professor of International Relations at Nottingham University, who questioned whether 'international recognition is ... the self-evident proposition it is sometimes made to seem. International excellence is properly *comparative*. Is the RAE', he asked, 'well designed to make such judgments?'<sup>77</sup> David Eastwood (who sat on the Roberts review) went further, conceding that 'international benchmarking of quality' was 'one thing that the RAE has not been able to do' – which might be thought a trifle odd, given that the scale used to evaluate each and every research output in RAE 2008 claimed to distinguish degrees of international excellence. Eastwood still had hopes then that the REF would solve the problem with 'bibliometrics, used with sensitivity and sophistication',<sup>78</sup> but we know where that proposal ended up. The only acknowledgement in REF 2014 that appraisal 'with reference to international quality standards' might benefit from involving the occasional foreigner is the presence of international members on Main Panels, who are supposed to ensure that 'subpanels adhere to internationally referenced standards'.<sup>79</sup> They are in a small minority, especially in the social sciences (4/22) and humanities (4/20), and it is hard to see how significant an input or effective an oversight they can have.<sup>80</sup> Like much else in the REF their main function appears to be window-dressing.

One might, I suppose, attempt to justify this idiosyncrasy by pointing to the UK's disproportionate international excellence in research. The

<sup>76</sup> *Review of research assessment: Report by Sir Gareth Roberts*, 6. Emphasis added.

<sup>77</sup> Alex Danchev, 'Goodbye to the RAE ... and hello to the REF'. *Times Higher Education*, 30 November 2007. Emphasis added.

<sup>78</sup> David Eastwood, 'Goodbye to the RAE ... and hello to the REF'. *Times Higher Education*, 30 November 2007.

<sup>79</sup> REF 2014: *Units of Assessment*, para 33.

<sup>80</sup> Here and below I draw on REF 2014: *Panel Membership Lists (updated December 2013)*. <http://www.ref.ac.uk/panels/panelmembership/> (accessed 13 January 2013).

Department of Business, Innovation and Skills (DBIS), under whose aegis UK universities currently fall, boasts that ‘while the UK represents just 0.9% of global population, 3.2% of R&D expenditure, and 4.1% of researchers, it accounts for 9.5% of downloads, 11.6% of citations and 15.9% of the world’s most highly-cited articles’.<sup>81</sup> Like the Russell Group in other contexts, DBIS has no qualms about using citation data as an indicator of the ‘international comparative performance of the UK research base’, maintaining that ‘Normalised citation impact *does* correlate with other assessments of research quality ... and so it can be seen as a good indicator of research performance – especially for larger samples’.<sup>82</sup> But – to give these numbers a somewhat different spin – while 95.9% of researchers, 90.5% of downloads, 88.4% of citations and 84.1% of the world’s most highly-cited articles do *not* hail from Britain’s universities, almost 100% of REF evaluators do. The academic members and assessors on the subpanels – *the people who actually do all the reading and scoring of outputs* – work in UK institutions unless they moved abroad after they were appointed. Graeme Rosenberg responded to criticisms that a subpanel chair who had moved to Sciences-Po in Paris after appointment would not be able to do his job properly by observing: ‘We have a number of people on the panels from overseas or with overseas experience or interests. This is a distinct advantage as the panels are, after all, judging work against international standards’.<sup>83</sup> Quite.

I imagine that it was costs and logistics rather than post-imperial hubris that led HEFCE to assume that assessors drawn from UK universities alone would suffice to evaluate close to 200,000 individual research outputs against international standards, but the presumption – in all senses – is gargantuan. No doubt it says something about British elites’ perceptions of their own superiority. Not only does REF practice in this respect starkly conflict with peer review in academic journal and book publishing, where reviewers are sought on the basis of their knowledge of the field irrespective of their national affiliation, but it deviates equally sharply from peer review in research grant competitions, where funding agencies also use disciplinary panels to decide on awards. The key difference with the REF is that although members of such panels may be drawn from a single national university system, their judgements are informed by external appraisals from all over the world. Canada’s SSHRC – of which I have

<sup>81</sup> *International Comparative Performance of the UK Research Base – 2013*. A report prepared by Elsevier for the UK’s Department of Business, Innovation and Skills (BIS), 2.

<sup>82</sup> *International Comparative Performance of the UK Research Base*, 7. Emphasis added.

<sup>83</sup> Quoted in Paul Jump, ‘Paris post will not weaken panellist’s “duty” to REF’. *Times Higher Education*, 28 March 2013.

had experience as an applicant, a three-time grant-holder, an expert reviewer and a panellist<sup>84</sup> – uses ‘appropriate expert reviewers to assess the individual merit of all proposals against the published criteria’ who ‘are enlisted based on individual experience and expertise, and ... may be from Canada or abroad’.<sup>85</sup> But it *separates* ‘the merit review of proposals ... from the making of funding decisions. Those acting as reviewers will not also be responsible for authorizing the funding decision’, which is made by disciplinary panels that score and rank all applications.<sup>86</sup> In tenure and promotion proceedings – including in many British universities – it is often obligatory for a file to be sent to international referees, especially at senior levels.

This is not the only problem with the composition of REF panels. We saw earlier that reducing RAE 2008’s 67 subpanels to REF 2014’s 36 led some to question their ability to provide adequate coverage of the fields that came under their remit. The slimming down of the panels merely highlights a more general problem. No subpanel small enough to be capable of working effectively is likely to contain the necessary ‘breadth and depth of expertise’ to provide equally informed assessment across all the sub-fields of any academic discipline (let alone interdisciplinary work, as REF 2014 now requires); and this inequity in turn casts doubts on the REF’s ability to deliver judgements that are either competent or fair. Let me take the history panel, whose remit has not changed since RAE 2008, as an example. The panel has a chair, a deputy chair and 23 members, of whom three are ‘user members’ representing the Wellcome Trust, Minerva Media and Historic Royal Palaces.<sup>87</sup> There are a further 14 assessors, 9 of whom are research

<sup>84</sup> I have held SSHRC grants from the History, Sociology and Interdisciplinary Studies committees and acted as a reviewer for all three. I served as a member of SSHRC’s disciplinary panel for anthropology, sociology and cognate disciplines in the 2013–2014 postdoctoral fellowships competition.

<sup>85</sup> SSHRC, ‘Merit Review’. [http://www.sshrc-crsh.gc.ca/funding-financement/merit\\_review-evaluation\\_du\\_merite/index-eng.aspx](http://www.sshrc-crsh.gc.ca/funding-financement/merit_review-evaluation_du_merite/index-eng.aspx) (accessed 18 January 2014).

<sup>86</sup> SSHRC, ‘Principles for Merit Review’. [http://www.sshrc-crsh.gc.ca/funding-financement/merit\\_review-evaluation\\_du\\_merite/index-eng.aspx](http://www.sshrc-crsh.gc.ca/funding-financement/merit_review-evaluation_du_merite/index-eng.aspx) (accessed 18 January 2014). More than one panel member reads each proposal, and where there is a significant discrepancy between scores, the case will be discussed and, if necessary, voted on by the entire panel. See *SSHRC Manual For Adjudication Committee Members 2013–2014*. [http://www.sshrc-crsh.gc.ca/funding-financement/merit\\_review-evaluation\\_du\\_merite/adjudication\\_manual\\_2013-guide\\_membres\\_2013-eng.pdf](http://www.sshrc-crsh.gc.ca/funding-financement/merit_review-evaluation_du_merite/adjudication_manual_2013-guide_membres_2013-eng.pdf) (accessed 18 January 2014).

<sup>87</sup> ‘Main Panel D Membership (December 2013)’. <http://www.ref.ac.uk/panels/panellistmembership/> (accessed 13 January 2013).

users (and therefore not eligible to evaluate outputs).<sup>88</sup> The 27 academic members and assessors on the panel have to read close to 7000 outputs from over 1750 researchers, potentially covering all periods of history and areas of the world.<sup>89</sup>

Attempts have clearly been made to achieve various kinds of balance, though it is not always an especially equitable one. There are 16 men and 11 women<sup>90</sup> among the evaluators on a panel that (so far as I can tell) is uniformly white. Universities from all parts of the UK are represented, with only four having more than one panel member or assessor and none more than two. Nineteen evaluators hail from Russell Group universities, a further six from the 1994 Group and only two from institutions outside the magic circle (the Universities of Keele and Hertfordshire). The coverage of medieval (3), early modern (5)<sup>91</sup> and modern history (19) is probably a fair reflection of the chronological distribution of the outputs to be assessed. This may also be the rationale for the skewed – and decidedly limited – geographic range of the panel's expertise. Twelve panellists and four assessors are historians of Britain;<sup>92</sup> of these, three also work in imperial history.<sup>93</sup> There are six historians of various European countries, two historians of the United States and one of Africa (which thus gets the same coverage as Wales, which may or may not say something about the

<sup>88</sup> The five academic assessors were added in June 2013, probably in the interest of covering gaps in coverage or relieving workload. I exclude user member/assessors in the calculations that follow, since my major concern is with the panel's range of academic expertise.

<sup>89</sup> Based on figures for 2008 RAE submissions given in *REF 2014: Units of Assessment*, Annex B, 13. The actual figure for outputs might be lower in practice due to double weighting of books, but this scarcely eases the task of evaluation since in most cases the length of a history book amounts to many times more than twice that of an article.

<sup>90</sup> Sharon Montieth (Nottingham University) is jointly appointed to Sub-panel 29, English Language and Literature.

<sup>91</sup> I include in the early modern count one intellectual historian (Colin Kidd, St Andrew's) whose interests span the periods historians distinguish as 'early modern' (16th–18th centuries) and 'modern' (later 18th century onwards). In characterizing panellists' research expertise I have drawn on their pages on university websites.

<sup>92</sup> I include here Chris Williams (Cardiff), a historian of Wales, and Keith Jeffery (Queen's University, Belfast), some of whose work deals with Ireland (at a time when it was ruled from Britain).

<sup>93</sup> Margot Finn (UCL; India), Clare Anderson (Leicester; India/South Asia) and Keith Jeffery.

horizons of history as a discipline in the UK).<sup>94</sup> Two panellists define their expertise thematically rather than geographically.<sup>95</sup> Nobody appears to have specialist knowledge of Spain, Portugal, Central and Eastern Europe (apart from Russia), the Middle East, Latin America, Australasia, China, Japan or anywhere else in Asia except where that continent's long and eventful history had the good or bad fortune to intersect with that of imperial Britain. There is a decent spread of sub-fields across economic, social, political, cultural, military, religious and intellectual history, while non-traditional areas such as gender, consumption, emotions and the body are also represented.

Whether this hodgepodge amounts to 'an *appropriate* breadth of research expertise' to judge the originality, significance and rigour of *every* submitted history output from *every* university in the UK – which is what, after all, the panel is supposed to provide – is questionable. I myself rather doubt it, but then I am not the 'community' whose confidence HEFCE claims REF panels command. What is *not* debatable is that the chances of outputs in history being read by panellists who are experts in an author's field are very unevenly distributed. Structuring the panel to give the evaluators comparable workloads may make organizational sense but it creates serious inequities for those whose research is being assessed. The odds of a modern historian's outputs finding their way to an evaluator who is familiar not only with his or her period, but also with the same country and maybe even the same research area are a good deal higher than for early modernists or medievalists simply because there is a wider spread of modernists on the panel. For the same reason historians of Britain are much more likely to find a reader who works on the same period or in the same field than historians of Germany, France, Russia or Holland (each of which have one representative on the panel). The REF may claim to assess 'all types of research and all forms of research output ... on a fair and equal basis',<sup>96</sup> but in this context some histories are more equal than others.

<sup>94</sup> Paul Nugent (Edinburgh). Though Nugent has pan-African interests and is joint editor of the *Journal of Modern African Studies* and president at AEGIS – African Studies in Europe, his personal expertise is Ghana, Togo, Senegal, Gambia and South Africa. I imagine asking him to evaluate an article on, say, Egypt, would make as much sense as asking a Swedish historian to appraise an article on Montenegro. See <http://uk.linkedin.com/pub/paul-nugent/21/8b9/62b> (accessed 13 January 2014).

<sup>95</sup> Catherine Schenk (Glasgow), an economic historian who has done work on Hong Kong; John Young (Nottingham), a historian of foreign policy who has worked on the UK, US and EU.

<sup>96</sup> REF 2014: *Panel Criteria*, para 40. Emphasis added.

Let us suppose that a paper titled 'Religious Sects in Seventeenth-Century Russia: A Revisionist Interpretation' is submitted to the History panel. Who should read it: the panel's lone Russian historian (whose period is the 20th century and whose field is political history)? One of its five early modernists (none of whom specialize in either Russian history or religion)? Its only ecclesiastical historian (whose area of research is early medieval England)? None are ideal readers, because none are experts *in the field to which the paper contributes*. I would insist that none – to return to the REF assessment criteria – are qualified to assess its *originality* or *significance*, since such judgements by definition demand knowledge of the current state of that field. None would figure on any journal editor's list of appropriate reviewers for our hypothetical article. Nor would any be likely to be invited to provide specialist assessments of a proposal on such a topic in grant competitions (though they might well serve on an AHRC or ESRC adjudicating panel), or be nominated as suitable external referees in a tenure or promotion case for an author in this field. This is not because they lack eminence in the historical profession, but because they do not have – in HEFCE's words – 'the expertise to reach robust and fair judgments with regard to the submitted material'.<sup>97</sup> Professional eminence and specialist expertise are two different things, and the one does not always entail the other.

Should all papers on American history, then, be divided up between the panel's two US historians? To do so at least guarantees a reading by somebody with expertise on the country, if not necessarily on the period and not likely in the field. But it also raises the spectre of gatekeeping. Does the academic community *really* want all outputs submitted to the REF on American history to be judged by just two people? Taking the point further, should everything on Russian history be appraised by the one Russian specialist on the panel, everything on Germany by the one German specialist – and everything on the entire continent of Africa evaluated by the one African historian? To do so is to grant these individuals' opinions extraordinary weight in a context in which the financial and reputational stakes are unusually high. Again, standard peer review procedures are designed to avoid this dilemma by soliciting multiple appraisals. But the only alternative within the REF framework is to trust other members of the panel to make the fine distinctions between 2\*, 3\* and 4\* contributions to fields that are not their specialties. The *reductio ad absurdum* comes with the history of China, Turkey or Brazil, or – as in the case of my own research – the Czech Lands, where there is nobody on the

<sup>97</sup> REF 2014: *Panel Criteria*, para 31. Emphasis added.

panel 'with appropriate expertise' (or possibly any knowledge at all). They don't know the languages, they don't know the archives, they don't know the sources and they don't know the secondary literature. On what possible basis can they determine whether an output in these fields is 'internationally excellent' or merely 'internationally recognized' – the crucial boundary between 3\* and 2\* – especially when they are expressly forbidden to use any bibliometric or other contextual data? Whatever the panel decides to do in such cases, by no stretch of the imagination can it be described as expert review.

History may be extreme in its diversity of chronological range, geographic span, substantive foci and methodological orientations (which may differ in their conceptions of 'rigour') but comparable issues must arise to a greater or lesser extent within every disciplinary subpanel. Earlier I quoted Peter Coles's doubts that the physics panel has sufficient 'breadth of understanding to do an in-depth assessment of every paper'. The underlying problem here is not the panels' composition per se. Composition becomes a problem only because the REF process *does all evaluation of outputs in-house*, refusing to solicit any external specialist advice; in some panels, *relies on a single panel member to evaluate each output*; and *disregards any and all external indicators*, including venue of publication, citations and book reviews. I am not denying the problems with bibliometric data, which opponents of metrics have rehearsed ad nauseam. But wilfully ignoring all consideration of whether an output has gone through a prior process of peer review, where it has been published, how it has been received and how often and by whom it has been cited in favour of the subjective opinion of just one evaluator who may have no expertise in the field does not seem to me a very defensible alternative. It also gives extraordinary gatekeeping power to the individuals on the REF subpanels.

Any one of these factors would seriously undermine the credibility of appraisals in any of the peer reviewing contexts – publication, grant funding, tenure and promotion decisions – examined in Chapter 1. Taken together they devastate it. The feature of the REF in which HEFCE takes most pride – its 'expert review of the outputs' – turns out to be exactly what its procedures are incapable of delivering in any reliable or consistent way. The truth of the matter is that the panels are a crapshoot and the appraisals a farce. Stevan Harnad and I may have had our disagreements over Open Access publication, but he went to the heart of what is wrong with the UK's research assessment system when (writing of the 2008 RAE) he stated bluntly '*that was not "peer review" in any case; peer review is done by journal referees, selected among the world's top experts, in the case of the best*



journals, not from a single country's rag-tag generic panel'.<sup>98</sup> Small wonder that every research university in the land was desperate to get its people on to REF panels.

## 2.4 THE RUSSELL GROUP, QR FUNDING AND INSTITUTIONAL GAMING

The precise funding formula that determines how the scores awarded for each UOA in REF 2014 will translate into cash for individual universities over the next five or six years is yet to be announced, but the general principles are clear. Government channels its support for university research through the 'dual system'. The larger share of the budget goes to the research councils, which annually invest around £3bn across all academic disciplines.<sup>99</sup> These investments mainly take the form of research grants and PhD studentships and need not concern us here. HEFCE and the other funding councils provide 'block grant funding for institutions to support the research infrastructure and enable ground-breaking research in keeping with their own mission'.<sup>100</sup> This is the QR funding discussed earlier. Unlike research council funding, which is tied to projects and studentships, universities can spend QR income as they see fit. It is allocated on the basis of *volume* of research (measured by the number of research-active staff entered in the RAE/REF), *costs* of research (laboratory-based subjects have a higher weighting) and *quality* of research as measured by RAE/REF performance. RAE 2008 gave each UOA a 'profile' showing the proportion of outputs that fell into each of the starred quality bands rather than an aggregate score as previously.<sup>101</sup> The relation between these profiles and the income a university would receive in HEFCE research support was made crystal clear in the QR funding formula announced in January 2009, which weighted outputs on a scale

<sup>98</sup> Stevan Harnad, 'Citations ideas 1'. *Times Higher Education*, 4 January 2008. For my disagreements with Professor Harnad on Open Access see <http://coastsofbohemia.com/2013/03/04/more-on-open-access-hefce-brings-out-the-big-ref-stick/> and <http://coastsofbohemia.com/2013/03/07/april-fools-day-rcuk-adds-fuel-to-the-open-access-fire/> (posted 4 and 7 March 2013, respectively).

<sup>99</sup> Research Council UK (RCUK) home page. <http://www.rcuk.ac.uk/Pages/Home.aspx> (accessed 16 January 2013).

<sup>100</sup> HEFCE, 'How we fund research'. <http://www.hefce.ac.uk/whatwedo/rsrch/howfundr/> (accessed 16 January 2014).

<sup>101</sup> HEFCE, *Securing world-class research in UK universities: Exploring the impact of block grant funding* [nd], 17. <http://www.hefce.ac.uk/research/funding/resfund/QR.pdf> <http://www.universitiesuk.ac.uk/highereducation/Pages/WorldClassResearch.aspx#U1VUm8bwDDE>