

# The Research Excellence Framework and the ‘impact agenda’: are we creating a Frankenstein monster?

Ben R Martin

In pursuit of public accountability, the mechanisms for assessing research performance have become more complicated and burdensome. In the United Kingdom, the Research Assessment Exercise (RAE) evolved from an initially simple framework to something much more complex and onerous. As the RAE now gives way to the Research Excellence Framework (REF), ‘impact assessment’ is being added to the process. Impact comes in numerous forms, however, so its assessment is far from straightforward. While the Higher Education Funding Council for England is initially proposing a relatively simple methodology, the history of the RAE suggests that this approach will over time become ever more sophisticated. Yet if the ‘costs’ of an elaborate system for assessing ‘research excellence’ and its impact then exceed the benefits, the time may have come to re-examine whether a dual-support system still represents the optimum way of funding university research.

**D**URING THE LAST 20 to 30 years, publicly funded research has become subject to ever more intensive accountability. Various forms of research assessment have been developed, each more complicated, burdensome and intrusive than its predecessor. In the UK, the process began during the 1980s, when Margaret Thatcher and her government required that all areas of public expenditure should demonstrate ‘value for money’ and provide evidence that the funds spent met the ‘three Es’ of economy, efficiency and effectiveness (e.g. Rhodes, 1994: 144). With regard to government spending on universities through the University Grants Committee (which was subsequently replaced by the University Funding Council, and later by the Higher Education Funding Council for England, HEFCE),<sup>1</sup> this gave rise in 1986 to the first in a series of periodic Research Assessment Exercises (RAEs).<sup>2</sup>

Over the following years, the UK research councils were also required to ensure that the research they funded was more closely related to the needs of ‘users’ in industry and elsewhere.<sup>3</sup> More recently, they have begun to request that research proposals should explain how the planned research is likely to have an impact and to outline the ‘pathways’ by which that impact is likely to be achieved.<sup>4</sup> Moreover, HEFCE has announced that the RAE is to be replaced by a new Research Excellence Framework (REF), in which assessment of the impact of previously funded research figures prominently.<sup>5</sup> What are likely to be the longer-term effects of the new REF and of implementing the ‘impact agenda’?

As publicly funded researchers, we clearly have a responsibility to contribute something in return to society — in other words, there is some form of a social contract that we are obliged to honour. As part of this, there is an obvious need for public accountability. As tax-payers (with a rather large proportion of our income going on income tax, national insurance, value added tax, council tax, fuel tax and so on), we, too, want to know that ‘our’ money is well spent. Moreover, we must accept that the need for accountability extends to all areas of government funding, including the funding of academic research.

---

Ben R Martin is Professor of Science and Technology Policy Studies at SPRU – Science and Technology Policy Research, The Freeman Centre, University of Sussex, Brighton, BN1 9QE, UK; Email: B.Martin@sussex.ac.uk; Tel: +44 1273 873562.

For acknowledgements see page 252.

At the same time, however, there are certain fundamental boundary conditions that surely need to be attached to the extent of this accountability. First, the costs of the assessment (where the 'costs' are broadly defined to include all the negative consequences and other indirect costs as well as the direct costs) needed to provide such accountability should never represent more than a very small proportion of the costs of the activities being assessed (e.g. less than 1% — Irvine and Martin, 1984). Second, and more specifically, the total costs should clearly be less than the benefits that flow from that enhanced accountability. Yet these fairly obvious boundary conditions seem to be in danger of being overlooked in more recent efforts to assess research and its impact.

Over the last 20 years, the methods put in place to assess research have become more sophisticated, partly because earlier efforts were criticised for being too crude or too 'unfair', but also thanks to the efforts of our own research evaluation community in improving the range of evaluation tools available. More sophisticated research assessment methods are almost inevitably also more costly and burdensome. In the UK and elsewhere, we are now adding 'impact assessment' to the earlier research assessment. Yet, as the Brunel workshop reported in this special issue of *Research Evaluation* revealed, 'impact' is extremely complicated to assess. Papers presented by workshop participants demonstrated that, with very considerable effort, it may be possible to assess the impact of research in a small number of case studies with quite a sophisticated methodology. However, such methodologies are particularly labour-intensive and are very much a 'craft activity' requiring extensive professional expertise (Sennett, 2008). Even so, there are major conceptual problems (What precisely is 'impact'?) and methodological problems (How can we assess it in a rigorous, reliable and comparable manner?) to be overcome.

Yet funding agencies like HEFCE are now proposing to convert this craft activity of impact assessment into mass production almost overnight. To do so, they will not be able to adopt a labour-intensive approach such as the Research Payback Framework (Klautzer *et al*, 2011; Nason *et al*, 2011; Scott *et al*, 2011) or the SIAMPI model (Spaapen and van Drooge, 2011; Molas-Gallart and Tang, 2011). Instead, by necessity, they will have to adopt something cruder. That crude approach will be even

more vulnerable to the sorts of criticisms voiced at the workshop by John Brewer<sup>6</sup> and others. As a result of such criticisms, and armed with expert advice from the impact assessment research community, the REF will over time doubtless become more sophisticated but also more burdensome. In short, we may be in danger of creating a Frankenstein monster. Furthermore, in the absence of open debate about the dangers that lie ahead, we, the impact assessment research community, may become complicit participants in that process.

The remainder of this article sets out in more detail the reasons why one might fear this outcome, particularly with regard to the direction that REF and its impact component are likely to take.<sup>7</sup> The second section briefly examines the origins of research evaluation and the evolution of the UK RAE. The third section considers the questions of 'What is impact?', and 'How can one assess it?', while the fourth section explains why an element of balance is required in constructing impact indicators in order to keep the costs well below the benefits. The final section examines where we might go next. In particular, it raises the question of whether the cost, the burden and the unintended consequences of the proposed REF are so great that we have reached the stage where we now need to re-examine whether the dual-support system for funding university research still constitutes the optimum solution.

## Origins of research assessment

The origins of work on research evaluation and on assessing the impact of research on the economy and society can be traced back to the 1970s. At that stage, there was significant interest among scientists and others in the issue of 'science and society' (e.g. Rose and Rose, 1970) and in the social responsibility of the scientist (e.g. Rip and Boeker, 1975; Nelkin, 1979). Earlier work by Bernal (1939) and other radical scientists (see Werskey, 1971) had raised doubts as to whether science was being used for the full benefit of humanity. The 1970s also witnessed a succession of economic crises, resulting in severe public expenditure cuts. This increased the premium on ensuring that the limited funds available for research were spent as wisely as possible. However, at that stage, most decisions within science were made behind closed doors and on the basis of peer review alone. Partly because of the limitations of peer review, it was argued that one needed to open up the decision-making process, providing a more public form of accountability (Irvine and Martin, 1984). For that to be achieved, one required data on the inputs to and the outputs from research (and, in due course, on the outcomes and impact of research).

SPRU (then named the Science Policy Research Unit), along with other organisations such as CHI Research (e.g. Narin, 1976) in the United States and the CWTS group at Leiden University headed by

---

**Yet funding agencies like HEFCE are now proposing to convert this craft activity of impact assessment into mass production almost overnight**

---

Antony van Raan (e.g. Moed *et al.*, 1985), was among the pioneers in developing policy-related tools for research evaluation in the late 1970s and early 1980s (e.g. Martin and Irvine, 1981, 1983; Irvine and Martin, 1983). These early efforts encountered initial opposition from the scientific community, who were concerned about outsiders encroaching on 'their' territory and challenging the authority of peer review as the primary decision-making mechanism (Martin, 1997). Gradually, however, under political and economic pressure, funding bodies such as the UK research councils began to adopt more systematic approaches to research evaluation, and over time these became more extensive.

In 1986, the University Grants Committee (UGC), recognising that a more selective approach was required to the funding of British university research, launched the Research Selectivity Exercise. This was comparatively simple and not too time-consuming, with panels carrying out a peer-review assessment of short submissions from each university department (or more precisely each 'research unit') and that department's best five publications. However, fierce criticism from academics<sup>8</sup> forced UGC and its successor, the University Funding Council (UFC) to make the approach in the 1989 Research Assessment Exercise (RAE) more rigorous but also more labour-intensive in terms of both preparing each unit's submission but also with respect to carrying out the assessment. The pattern was to be repeated over each successive RAE, with criticisms of the previous exercise resulting in yet more refinements being introduced in the next one (e.g. Martin and Whitley, 2010). Hence over time, the RAE became more sophisticated and perhaps better suited to the difficult and complex task of assessing the research of each university department. At the same time, however, the RAE was proving ever more burdensome and intrusive (Hicks, 2011: 9–10), with various perverse incentives<sup>9</sup> being introduced that would distort academic research (Martin and Whitley, 2010).

As the pressures to do well in the RAE increased, so universities and their departments came to put growing levels of effort into preparing for the next exercise. If one includes all the time devoted to this preparatory work as well as the time spent by RAE panels assessing the detailed submissions from each university unit in their field, then the total costs of the RAE grew to become very considerable, perhaps of the order of £100 million (Sastry and Bekhradnia, 2006: 5).

At the same time, while the early RAEs undoubtedly brought significant benefits in the form of greater attention to published research and a more strategic approach to research by universities and departments, those benefits were almost certainly subject to diminishing returns. At a certain point, a graph of increasing costs must inevitably cross the graph of diminishing benefits, and Geuna and Martin (2003) argued that this had probably occurred

some time after about the third or fourth RAE (carried out in 1992 and 1996 respectively). This paper contributed to a debate in the mid-part of the first decade of the 21st century, with others also recognising that the RAE had become too costly and burdensome. (After all, an assessment costing of the order of £100 million, and involving the evaluation of some 50,000 individual researchers and nearly 200,000 publications, is a rather complicated and expensive way of determining how to divide up the government university research budget between 100 or so universities!)

The Treasury proposed that perhaps this peer-review-based assessment could be replaced simply by metrics such as the funding received from research councils and other funding bodies (HM Treasury, 2006). However, this brought fierce opposition from academics loath to see the disappearance of peer review from the assessment process. Instead, HEFCE and the other funding councils finally decided to replace the RAE with the Research Excellence Framework (REF), a combination of the previous peer-review-based exercise with some form of impact assessment.

### What is 'impact'? How can one assess it?

This brings us to the question of 'impact' — what precisely is it? No one is very sure. In general terms, impact is clearly important. As argued above, publicly funded researchers now operate under a fairly explicit social contract in which, in return for public funding, they are expected to contribute something in return to society. The problem is that each research community (and perhaps even each individual) has their own conception of impact as it pertains to their research. Impact as perceived, say, by an engineer is very different from that of the medical researcher or the sociologist or the historian. How can all these different conceptions of the nature of impact be reconciled in a single assessment mechanism?<sup>10</sup>

In addition, impact clearly comes in differing magnitudes — in a few instances the impact may be extremely large while in most cases it is much more modest. How can one come up with an appropriate 'scale' to assess and measure impacts of very different magnitudes? Another complicating factor is that not all impact may be desirable or 'good'. For example, Andrew Wakefield's research on an apparent link between vaccination and autism (Wakefield *et al.*, 1998)<sup>11</sup> clearly had an enormous social impact (in terms of falling vaccination rates), but in the view of many medical researchers and doctors that impact would now be judged to have been negative. But who would judge whether an impact had been positive or negative? And how would they do so?

In addition, as several participants at the Brunel workshop noted, there are numerous other problems with impact, rendering its assessment very difficult if not impossible,<sup>12</sup> including the fact that it is often

---

**Experiences with university spin-offs illustrate a central and inescapable problem with any assessment system, namely that once you measure a system, you irrevocably change it**

---

indirect, partial, opaque and long-term.<sup>13</sup> Indeed, the very concept of 'impact' and its assessment implies a rather linear model of how knowledge from an individual piece of research is subsequently taken up and used (Martin, 2007). Yet as scholars in the field of science policy research have demonstrated over the last 30 years, such a linear model rarely applies (Martin *et al.*, 1996; Salter and Martin, 2001; Martin and Tang, 2006; Martin, 2012). As a result, there is a major and perhaps insuperable problem in trying to 'capture' impact, and doing so in a rigorous, consistent and complete manner across the full range of sciences, social sciences and humanities, and across over 100 universities and other higher education institutions.

There have been a number of attempts to assess the impact of research in the past. Perhaps the most comprehensive effort was that carried out in Australia as part of the preparations for the proposed Research Quality Framework (RQF)<sup>14</sup> (Donovan, 2007, 2008). However, these plans for assessing the wider economic, social, environmental, and cultural benefits of research:

sent mixed messages and created confusion about what impact is, and how it is best measured, to the extent that this bold live experiment did not come to fruition. (Donovan, 2008: 47).<sup>15</sup>

A comprehensive review of this and other efforts to assess research impact can be found in Grant *et al.* (2009).

Among the measures that have been used to assess research impact are numbers of patents and university spin-offs. Experiences with university spin-offs illustrate a central and inescapable problem with any assessment system, namely that once you measure a system, you irrevocably change it. There are obvious parallels here with Heisenberg's uncertainty principle (Heisenberg, 1930); that is, the more accurately you attempt to measure one characteristic (the position of a particle), the less you know about another (its momentum). A further analogy is with the Hawthorne effect in management,<sup>16</sup> where the mere presence of researchers brings about a significant change in performance of the workers being observed, a form of psychological 'placebo effect' (Wickström and Bendix, 2000). There is also a similarity with Goodhart's Law, according to which,

once an economic indicator becomes a performance target for particular policy purposes, it rapidly loses its ability to serve as an appropriate and reliable indicator of the chosen characteristic (Goodhart, 1975; see also Chrystal and Mizen, 2001).

Around 2001, the British Science Minister, Lord Sainsbury, started to refer to university spin-offs as a useful indicator of technology transfer from universities. Over the previous five years, an average of 67 university spin-offs had been created, but within just a year the figure had leapt to 199 (Davis, 2002). This was taken to herald a sudden and dramatic increase in the level of university technology transfer, apparently by a factor of three. Great play was made of the fact that the 'cost per spin-off' (calculated as the annual government spending on university research divided by the number of spin-offs set up in a year) was now far lower than in the United States (Morgan, 2002). A more accurate picture of what was happening was that British universities, aware that this indicator was now being used for policy purposes, strongly encouraged their academics to launch start-up companies, yet most of these never amounted to much more than a registered company name and a single employee (the founder). In other words, the shelf-life of this particular performance indicator had been just a year or so, as the actors involved adjusted their behaviour in order to maximise their 'score'. If the level of knowledge transfer from British universities had increased at all, it certainly had not grown by a factor of three.

### **Balancing cost and sophistication in developing impact assessment**

As organisations set about the task of assessing impact, they have essentially two broad options. One is to construct a relatively simple assessment procedure based on one or two chosen characteristics or dimensions of impact, perhaps with a few relevant indicators. However, the inevitable consequence will be that academics, as intelligent rational actors, will then modify their behaviour to maximise their score or rating on these particular dimensions. As the example described above illustrates, this is not necessarily the same thing as achieving a larger or better impact. It may instead take the form of skewing their activities in a few defined directions captured by the chosen impact assessment methodology.

Alternatively, funding agencies may decide to develop a much more complex assessment system designed to capture impact in all its multifarious forms.<sup>17</sup> However, this would almost certainly require the compilation of a wide range of evidence and numerous indicators. For example, when Molas-Gallart *et al.* attempted to identify a full range of indicators to encapsulate the outcomes from 'third mission' activities at universities, they concluded that over 60 indicators would be required to tackle the task comprehensively (Molas-Gallart *et al.*, 2002:

Appendix 1). A comprehensive assessment of impact would almost certainly require a similar number of indicators. However, such a complex assessment system would be very costly to set up and to maintain (all the data required to operationalise the various indicators would need to be regularly updated). There would be not just the direct financial costs, but also the elaborate bureaucracy to maintain and operate the system.

This would then raise the question of whether all those costs were actually greater than the benefits brought about by such an assessment system. As noted earlier, that crossover point was probably reached in the late 1990s after three or four RAEs. For the new REF (i.e. essentially the old RAE but with impact assessment added to it), the benefits are unlikely to be much greater than those from the former RAE (where most of the 'easy' benefits were obtained in the earlier exercises), while the costs (both direct and indirect) are likely to be greater now that impact assessment has been added to it.<sup>18</sup>

Moreover, while HEFCE is currently proposing to start with a relatively simple process for assessing impact (with each department submitting one impact case study for every eight to 10 academics; see HEFCE, 2011b: 26–30), this will almost certainly be subject to the same sorts of criticisms as the RAE was in its original simple form (when departments were just required to submit their five best publications). The probable HEFCE response will be to make it more sophisticated with each successive version of the REF, following a similar trajectory to the RAE. The first REF, like the early RAEs, may indeed result in some benefits such as incentivising behaviour that strengthens the broader relevance of research and hence the science–society relationship (Nightingale and Scott, 2007). However, diminishing returns would then set in, probably quite quickly, while the costs of an ever more sophisticated assessment of impact would rise, resulting in a crossover of the cost and benefit curves before too long.

### Where next?

The most likely outcome (and the one on which we already seem to be embarked) is that universities and their faculty will reluctantly comply with the panopoly of impact assessment mechanisms being put in place by HEFCE and the research councils. As these are applied, there will be criticisms that these mechanisms fail to capture certain aspects of impact,<sup>19</sup> so additions will be made to the assessment machinery, adding to the costs and the compliance burden, encouraging more 'game-playing', introducing more perverse incentives and generating more unintended consequences. All this will bring further complaints and probably yet more 'refinements'. By then, senior university managers and heads of department will all be busy impressing upon their staff the importance of 'playing the game' to the best of their ability.

---

**Consequently, like the proverbial frog in a saucepan of water the temperature of which is being gradually raised, university researchers may be unable to summon the effort needed to jump out of this cauldron as it comes to the boil**

---

Consequently, like the proverbial frog in a saucepan of water the temperature of which is being gradually raised, university researchers may be unable to summon the effort needed to jump out of this cauldron as it comes to the boil (e.g. Tichy and Devanna, 1986: 44; Hambrick and D'Aveni, 1988: 4).

What are the alternatives? One would be for HEFCE, armed with evidence from the impact assessment research community, to tell the government clearly and firmly that, while it could in principle provide the required accountability through the REF, the full costs of this mechanism, once all the indirect costs and distorting effects are taken into account, are greater than the benefits. As such, the REF would then surely represent an inappropriate use of public funds and it should be dropped.

The other, more radical, alternative is to launch a debate about the pros and cons of the dual-support system of university research. The dual-support system has in the past brought many undoubted benefits. In particular, the institutional 'core' funding of universities has provided them with resources that they are free to allocate wherever they choose, enabling them to 'pump prime' new areas and new researchers as well as providing the basic research infrastructure of the 'well found' laboratory (ABRC, 1987: 5, 11, 33). However, since the 1980s HM Treasury has required that this 'first stream' of research funding (the second being that flowing from research councils and other funders for specific projects) be subject to public accountability — hence the creation of the RAE. Consequently, despite its considerable benefits, this stream of funding now comes with a very substantial 'cost', namely the RAE and more recently the REF.

In previous decades, it was generally assumed that the benefits of the dual-support system outweighed the costs. Hence, this system was adopted in most countries around the world, although, interestingly, not in the United States,<sup>20</sup> where universities receive core funding for teaching but not research, and where faculty must therefore apply to funding bodies for the full costs of their intended research.

Now, after a succession of RAEs and with the forthcoming REF, the balance between the benefits and costs of the dual-support system has undoubtedly shifted. Whether the benefits of the dual system

are still greater than the costs is no longer clear. What we therefore urgently need is a full and open debate about the pros and cons of the dual-support system, and the respective pros and cons of the alternative, namely a US research funding system in which university faculty would need to apply to the research councils (or other funders) for the *full* costs of their proposed research, rather than having a proportion of them (currently around 20%) being met from the institutional funding provided by HEFCE.<sup>21</sup>

Such a US-style system would not be without its disadvantages, including the need for academics to spend a considerable part of their time preparing numerous research proposals, many of which were unsuccessful. In addition, there is a danger that such a system might lead to a further concentration of resources on a small group of elite universities. However, one great advantage of a US-style system is that the assessment would then focus on research that is about to be done, in contrast to the RAE which focussed on research published up to six years ago (and perhaps carried out a year or two before that), while the REF is to consider the impact of research conducted up to 20 years earlier, long after many of the individuals concerned have moved on from that department, changed their research interests, retired or died.

For too long, universities as well as successive governments have seemed to regard the dual-support system as a 'sacred cow' (even though there is surprisingly little rigorous empirical evidence as to its advantages). The time is surely ripe now for a systematic comparison of the respective advantages and disadvantages of the two alternatives — a dual-support system and a fully funded (US-style) research system, in which the 'quality research' stream of income currently coming from HEFCE would instead be switched to the research councils,<sup>22</sup> thereby abolishing at a stroke the need for the REF. One cannot be certain what the outcome of this debate would be. However, that only makes the need for such a debate all the more crucial. Without it, the university researcher may be condemned to the fate of the 'boiled frog'.

## Acknowledgements

This paper began as a response to the presentation made by Claire Donovan (partly written up as Donovan, 2011) at the HERG Impact Workshop on 'State of the Art in Assessing Research Impact', held at Brunel University on 31 March – 1 April 2011. The author is grateful for the feedback from other participants at the workshop, which encouraged him to develop this into a fuller article, and to Claire Donovan, the editor of this special issue of *Research Evaluation*, and two anonymous referees for their reviews, which helped to strengthen the paper.

## Notes

1. There are similar bodies for Scotland, Wales and Northern Ireland; together with HEFCE, they are known collectively as the Higher Education Funding Councils. In what follows, we

- refer to HEFCE for simplicity, while recognising that these other councils are also fully involved in the RAE and REF.
2. In 1986, this was known as the 'Research Selectivity Exercise', but by 1989 the name had changed to the 'Research Assessment Exercise' (RAE), and this remained the title for the next 20 years.
3. This requirement was formally set out in the 1993 Government White Paper entitled *Realising Our Potential* (OST, 1993).
4. See e.g. the guidelines on 'Pathways to Impact' at <<http://www.rcuk.ac.uk/kei/impacts/Pages/home.aspx>>; <<http://www.epsrc.ac.uk/funding/apprev/preparing/Pages/economicimpact.aspx>>; and <<http://www.esrc.ac.uk/funding-and-guidance/tools-and-resources/impact-toolkit/what-how-and-why/pathways/index.aspx>>, last accessed 12 July 2011.
5. See <<http://www.hefce.ac.uk/research/ref/>>, last accessed on 12 July 2011; and HEFCE (2011a,b).
6. Some of these criticisms can be found in Brewer (2011a,b).
7. A critique focussing more on the likely effects of the implementation by the Research Councils of their 'pathways to impact' can be found in Holmwood (2011).
8. A summary of some of those criticisms can be found in Phillimore (1989).
9. For example, the RAE is believed to have encouraged researchers to engage in research likely to result in quick publication rather than longer-term research, to produce 'mainstream' research more easily published in leading disciplinary journals rather than more 'risky' research, and to collaborate with researchers in other institutions rather than with departmental colleagues (Nightingale and Scott, 2007: 546–547; Martin and Whitley, 2010).
10. Although REF assessment panels will be allowed some latitude in interpreting how impact should be interpreted in the case of their particular discipline, they will nevertheless all work with the same 'broad generic criteria of "reach and significance" of the impact or benefit' (HEFCE, 2011a: 5).
11. The paper was retracted by the editors of *The Lancet* 12 years later in 2010, but until then it could presumably still have been submitted to the REF for impact purposes.
12. For a discussion of some of the conceptual, methodological and other problems, see Brewer (2011a,b); Collins (2011) and Holmwood (2011).
13. In particular, the process by which knowledge generated by research in the social sciences and elsewhere has an impact is far more likely to be one of gradual 'knowledge creep' rather than a simple, direct and instantaneous causative effect (Weiss, 1980).
14. The plans for the Australian Research Quality Framework were set out in DEST (2006).
15. Instead, the RQF was replaced in 2009 by the Excellence in Research for Australia (ERA) assessment scheme, in which there is no explicit attempt to assess the wider impact of research. <<http://www.arc.gov.au/era/>>, last accessed on 21 July 2011.
16. For a critical re-examination of the Hawthorne effect, however, see Adair (1984).
17. Some might hope that there is a happy medium to be struck here between a simplistic and a very complicated assessment system – a 'Goldilocks' form of impact assessment that is 'just right'. However, there must be doubts about how stable such a position is likely to prove in the longer term. Criticisms from academics that it failed to do justice to certain forms or dimensions of impact are likely to drive it to ever greater complexity (cf. Hicks, 2011: 9). Moreover, even with such a middle-of-the-road approach, academics will still play games to maximise their scores.
18. As Bekhradnia (2009: 6–7) notes: 'the introduction of impact as a factor in the REF will almost certainly substantially increase that [administrative] burden' on universities.
19. Already, criticisms of the HEFCE impact assessment plans have started to appear in print, even though there are still two more years before the REF is due to take place; see e.g. Smith *et al* (2011).
20. The USA considered adopting a dual-support system in the 1930s, but this was successfully resisted by the elite (primarily East coast) institutions, which feared that under such a system too great a proportion of the funds would be spread among large state universities in the mid-West.
21. In the late 1990s, the UK Government's Chief Scientific Adviser, Sir Robert (later Lord) May, proposed that the UK should switch from the dual-support system to a US system

of fully costed research funding. However, at that point there was no research council for the arts and humanities, which would have left academics in these fields with no obvious source of research support. In addition, universities at that stage had yet to calculate the full economic costs of research projects, although that was to change a few years later at the behest of the research councils. With these two major obstacles to May's original proposals now removed, the possibility of making such a switch no longer appears so outlandish.

22. Given that the budget of the Arts and Humanities Research Councils is much smaller than those of the other research councils, one would need to redistribute the HEFCE 'quality research' money in proportion to the numbers of academics in different fields, rather than *pro rata* to the existing budgets of the research councils. Without this, the funding of arts and humanities would suffer appreciably.

## References

- ABRC, Advisory Board for the Research Councils 1987. *A Strategy for the Science Base: a Discussion Document*, produced by the Advisory Board for the Research Councils. London: HMSO.
- Adair, John G 1984. The Hawthorne effect: a reconsideration of the methodological effect. *Journal of Applied Psychology*, **69**(2), 334–345.
- Bekhradnia, Bahram 2009. *Proposals for the Research Excellence Framework: a Critique*. Oxford: Higher Education Policy Institute.
- Bernal, John D 1939. *The Social Function of Science*. London: Routledge.
- Brewer, John 2011a. Viewpoint: from public impact to public value. *Methodological Innovations Online*, **6**(1), 9–12.
- Brewer, John 2011b. The impact of impact. *Research Evaluation*, **20**(3), September, 255–256.
- Chrystal, K Alec and Paul D Mizen 2001. Goodhart's Law: its origins, meaning and implications for monetary policy. Prepared for the Festschrift in honour of Charles Goodhart held on 15–16 November 2001 at the Bank of England.
- Collins, Harry 2011. Measures, markets and information. *Methodological Innovations Online*, **6**(1), 3–6.
- Davis, Caroline 2002. Analysis: UK's spin-offs reap £3 billion reward. *Times Higher Education*, 25 October 2002. <<http://www.timeshighereducation.co.uk/story.asp?storyCode=172516&sectioncode=26>>, last accessed 12 July 2011.
- DEST, Department of Education, Science and Training 2006. *Research Quality Framework: Assessing the Quality and Impact of Research in Australia: the Recommended RQF*. Canberra: DEST.
- Donovan, Claire 2007. The qualitative future of research evaluation. *Science and Public Policy*, **34**(8), October, 585–597.
- Donovan, Claire 2008. The Australian Research Quality Framework: a live experiment in capturing the social, economic, environmental, and cultural returns of publicly funded research. In *Reforming the Evaluation of Research: New Directions for Evaluation*, eds C L S Coryn and M Scriven, **118**, 47–60.
- Donovan, Claire 2011. State of the art in assessing research impact: introduction to a special issue. *Research Evaluation*, **20**(3), September, 175–179.
- Geuna, Aldo and Ben R Martin 2003. University research evaluation and funding: an international comparison. *Minerva*, **41**, 277–304.
- Goodhart, Charles A E 1975. Monetary relationships: a view from Threadneedle Street. In *Papers in Monetary Economics*, Vol. 1. Reserve Bank of Australia.
- Grant, Jonathan, Philipp-Bastian Brutscher, Susan Ella Kirk, Linda Butler and Steven Wooding 2009. *Capturing Research Impacts: a Review of International Practice*. Report prepared for the Higher Education Funding Council for England. Cambridge: RAND Europe.
- Hambrick, Donald C and Richard A D'Aveni 1988. Large corporate failures as downward spirals. *Administrative Science Quarterly*, **33**, 1–23.
- HEFCE, Higher Education Funding Council for England 2011a. *Decisions on Assessing Research Impact*. Bristol: HEFCE.
- HEFCE, Higher Education Funding Council for England 2011b. *Assessment Framework and Guidance on Submissions*. Bristol: HEFCE.
- Heisenberg, Werner 1930. *The Physical Principles of the Quantum Theory*. Chicago: University of Chicago Press.
- Hicks, Diana 2011. *Performance-based University Research Funding Systems*. Atlanta: Georgia Institute of Technology.
- HM Treasury 2006. *Science and Innovation Investment Framework 2004–2014: Next Steps*. Discussion paper prepared by HM Treasury, Department of Trade and Industry, Department for Education and Skills, and Department of Health. London: HMSO.
- Holmwood, John 2011. The impact of 'impact' on UK social science. *Methodological Innovations Online*, **6**(1), 13–17.
- Irvine, John and Ben R Martin 1983. Assessing basic research: the case of the Isaac Newton Telescope. *Social Studies of Science*, **13**, 49–86.
- Irvine, John and Ben R Martin 1984. What direction for basic scientific research? In *Science and Technology Policy in the 1990s and Beyond*, eds M Gibbons, P Gummert and B M Udgarkar, pp. 67–98. Harlow: Longman.
- Klautzer, Lisa, Stephen Hanney, Eddy Nason, Jennifer Rubin, Jonathan Grant and Steven Wooding 2011. Assessing policy and practice impacts of social science research: the application of the Payback Framework to assess the Future of Work programme. *Research Evaluation*, **20**(3), September, 201–209.
- Martin, Ben R 1997. Evaluating science and scientists: factors affecting the acceptance of evaluation results. In *Evaluating Science and Scientists: an East-West Dialogue on Research Evaluation in Post-Communist Europe*, eds M S Frankel and J Cave, pp. 28–45. Budapest: Central European University Press.
- Martin, Ben R 2007. Assessing the impact of basic research on society and the economy. Invited presentation at the FWF-ESF International Conference on 'Science Impact: Rethinking the Impact of Basic Research on Society and the Economy', Vienna, 11 May 2007.
- Martin, Ben R 2012. The evolution of science policy and innovation studies. *Research Policy* (forthcoming).
- Martin, Ben R and John Irvine 1981. Internal criteria for scientific choice: an evaluation of the research performance of electron high-energy physics accelerators. *Minerva*, **XIX**, 408–432.
- Martin, Ben R and John Irvine 1983. Assessing basic research: some partial indicators of scientific progress in radio astronomy. *Research Policy*, **12**, 61–90.
- Martin, Ben R and Puay Tang 2006. *The Economic and Social Benefits of Publicly Funded Basic Research*. Report to the Office of Science and Innovation, Department of Trade and Industry. Brighton: SPRU.
- Martin, Ben R and Richard Whitley 2010. The UK Research Assessment Exercise: a case of regulatory capture? In *Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences and their Consequences for Intellectual Innovation*, eds R Whitley, J Gläser and L Engwall, pp. 51–80. Oxford: Oxford University Press.
- Martin, Ben R, Ammon Salter, Diana Hicks, Keith Pavitt, Jacqueline Senker, Margaret Sharp and Nick von Tunzelmann 1996. *The Relationship between Publicly Funded Basic Research and Economic Performance: a SPRU Review*. London: HM Treasury.
- Moed, H F, W J M Burger, J G Frankfort and A F J van Raan 1985. The use of bibliometric data for the measurement of university research. *Research Policy*, **14**(3), 131–149.
- Molas-Gallart, Jordi and Puay Tang 2011. Tracing 'productive interactions' to identify social impacts: an example from the social sciences. *Research Evaluation*, **20**(3), September, 219–226.
- Molas-Gallart, Jordi, Ammon Salter, Pari Patel, Alister Scott and Xavier Duran 2002. *Measuring Third Stream Activities*. Brighton, UK: SPRU, University of Sussex.
- Morgan, Oliver 2002. Master of science fiction. *The Observer*, 20 October 2002. <<http://www.guardian.co.uk/business/2002/oct/20/interviews.politics>> on 12 July 2011, last accessed 12 July 2011.
- Narin, Francis 1976. *Evaluative Bibliometrics*. Cherry Hill, NJ: Computer Horizons Inc.
- Nason, Eddy, Brendan Curran, Stephen Hanney, Barbara Janta, Gillian Hastings, Mairead O'Driscoll and Steven Wooding 2011. Assessing the impacts of the Health Research Board of Ireland's funding activities. *Research Evaluation*, **20**(3), September, 193–200.
- Nelkin, Dorothy 1979. The social responsibility of scientists. *Annals of the New York Academy of Sciences*, **334**, 176–182.

- Nightingale, Paul and Alister Scott 2007. Peer review and the relevance gap: ten suggestions for policy-makers. *Science and Public Policy*, **34**(8), October, 543–553.
- OST, Office of Science and Technology 1993. *Realising Our Potential: a Strategy for Science, Engineering and Technology*. White Paper prepared by the Office of Science and Technology (Cmnd. 2250). London: HMSO.
- Phillimore, A John 1989. University research performance indicators in practice: the University Grants Committee's evaluation of British universities, 1985–86. *Research Policy*, **18**, 255–271.
- Rhodes, R A W 1994. The hollowing out of the state: the changing nature of the public service in Britain. *Political Quarterly*, **65**(2), 138–151.
- Rip, Arie and Egbert Boeker 1975. Scientists and social responsibility in the Netherlands. *Social Studies of Science*, **5**(4), 457–484.
- Rose, Hilary and Steven Rose 1970. *Science and Society*. Harmondsworth: Penguin.
- Salter, Ammon J and Ben R Martin 2001. The economic benefits of publicly funded basic research: a critical review. *Research Policy*, **30**, 509–532.
- Sastry, Tom and Bahram Bekhradnia 2006. *Using Metrics to Allocate Research Funds*. Oxford: Higher Education Policy Institute.
- Scott, Jack E, Margaret Blasinsky, Mary Dufour, Rachel Mandal and Stephane Philogene 2011. Assessing the impact of an NIH program using the Research Payback Framework. *Research Evaluation*, **20**(3), September, 185–192.
- Sennett, Richard 2008. *The Craftsman*. New Haven: Yale University Press.
- Smith, Simon, Vicky Ward and Allan House 2011. 'Impact' in the proposals for the UK's Research Excellence Framework: shifting the boundaries of academic autonomy. *Research Policy* (in press).
- Spaapen, Jack and Leonie van Drooge 2011. The role of productive interactions in the assessment of social impact of research. *Research Evaluation*, **20**(3), September, 211–218.
- Tichy, Noel M and Mary Anne Devanna 1986. *The Transformational Leader*. New York: Wiley.
- Wakefield, Andrew J *et al* 1998. Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *The Lancet*, **351**(9103), 637–641 (subsequently retracted).
- Weiss, Carol H 1980. Knowledge creep and decision accretion. *Knowledge: Creation, Diffusion, Utilization*, **1**(3), 381–404.
- Werskey, P Gary 1971. British scientists and 'outsider' politics, 1931–1945. *Science Studies*, **1**(1), 67–83.
- Wickström, G and T Bendix 2000. The Hawthorne effect: what did the original Hawthorne studies actually show? *Scandinavian Journal of Work, Environment and Health*, **26**(4), 363–367.