

Democratising research evaluation: Achieving greater public engagement with bibliometrics-informed peer review

Gemma E. Derrick^{1,*} and Vincenzo Pavone¹

¹CSIC Institute of Policies and Public Good, Madrid, C/Albasanz 26–8, Madrid, Spain 28037.

*Corresponding Author. Email: gemma.derrick@csic.es.

The ability of metrics to represent complex information about research in an accessible format has previously been overlooked in preference to debate about their shortcomings as research evaluation tools. Here, we argue that bibliometrics have the potential to widen scientific participation by allowing non-academic stakeholders to access scientific decision making, thereby increasing the democratisation of science. Government policies from 3 countries (UK, Australia and Spain) are reviewed. Each country outlines a commitment to the democratisation of science for one set of policies whilst ignoring this commitment when developing parallel research evaluation policies. We propose a change in dialogue from *whether* bibliometrics should be used to *how* they should be used in future evaluations. Future research policies should take advantage of bibliometrics to foster greater democratisation of research to create more socially-reflexive evaluation systems.

Keywords: democratisation of science; research evaluation; science and innovation policy; public engagement.

1. Introduction

As research moves towards more participatory approaches (Jasanoff 2004; Lengwiler 2008; Saurugger 2010), so too should research evaluation become more transparent and participatory. This participation should not only promote scientific excellence within the academic community but also improve accessibility for both the lay public and other members of the academic community. Considerable debate has been triggered by this participatory turn and while there is an emerging consensus of the need to improve scientific participation and its accountability, there is still little consensus on how to achieve these goals in terms of research policy (Rowe and Frewer 2005; Rowe et al. 2008; Felt et al. 2010).

Some authors highlight the advantage of bibliometrics as increasing transparency and accountability (Butler 2008) but none have yet considered these advantages in relation to the democratisation of research evaluation for the non-academic community, in line with a process

defined as the ‘democratisation of science’ (Liberatore and Funtowicz 2003). Irrespective of their academic controversy, the inherent utility of bibliometrics for translating bottom line information to non-academic stakeholders is frequently overlooked in favour of continued debate about their appropriate use and unsuitability in some research fields. As a result, the critics of metrics tend to overemphasise the merits of the traditional peer-review process, which is considered more suitable for preserving the autonomy of scientific self-government (Taylor 2011). Although it remains a fundamental element of research evaluation, traditional research evaluation via peer review fails to address the transparency and accountability of science. In fact, when used as the only method to perform research evaluation, it tends to preserve the process of research evaluation within the boundaries of a narrow epistemic community promoting a long criticised ivory tower approach. As such, it defies current trends towards a more open and democratic science.

Without ignoring their inherent limitations as evaluation tools, this paper supports a role for bibliometrics in research evaluation. When used in combination with appropriate peer-review methods, bibliometrics, as part of an informed peer-review process, has the potential to widen scientific participation by allowing non-academic stakeholders to access scientific decision-making, thereby making a significant contribution towards the democratisation of science and promoting a yearned-for model of stakeholder evaluation for research. In addition, and more importantly for the purpose of this paper, parallel government policies currently contradict each other by emphasising the value of the greater democratisation of science for one set of policies while ignoring this commitment when advancing research evaluation policies.

2. Democratisation of science and upstream participation in research

Participatory approaches to both science evaluation and policy regulation have been proposed mainly in order to: prevent technocracy; improve democratic accountability; encourage dialogue between scientists, policy-makers and the public; and secure strong and democratic support for science and innovation systems (Armstrong 2002; Felt and Wynne 2007). Yet governments increasingly rely solely on scientific expertise in order to deliver good regulation (Weingart 1999). When decisions are based on 'sound science', they are taken to be authoritative and acquire legitimacy (Eriksen et al. 2003; Nowotny 2003).

This issue is, however, controversial. Some authors argue that scientific judgment promises to bring neutrality to decision-making whereas politicisation leaves the outcomes in the hands of those who can exert greater decisional power (Kerr et al. 1997). In contrast, others have remarked that government-appointed experts often shape political decisions under the guise of 'scientific advice' (Weingart 1999; Todt and Luján 2000). As a result, various participatory schemes have been proposed and implemented for decision-making about science and technology innovation (Rowe et al. 2008). Often inspired by a 'democratisation of expertise approach', these schemes defend the potential of enacting new cooperative formats of mutual learning between scientists and stakeholders. This contributes to a wider definition of the issues as well as the formulation of viable solutions (Liberatore and Funtowicz 2003; Funtowicz and Ravetz 1993).

The 'democratisation of science' approach was first elaborated by Funtowicz and Ravetz (1993) in relation to what they defined as 'post-normal science'. The latter, as they argue, is a new area of science, mainly concerned with emerging biotechnologies, where high stakes are associated with a high level of scientific uncertainty. Post-normal science is therefore an area of science and innovation that cannot be left to traditional risk

assessment and expert-based decision-making, because the safety implications and potential harmful consequences are often unknown or extremely difficult to predict. As a result, Funtowicz and Ravetz suggest that science and policy-making decisions about research need to be democratised. Generally, people and stakeholders should be involved in the assessment and the decision-making process, so when decisions are made, they are the result of a larger social consensus based on shared values and priorities (Liberatore and Funtowicz 2003).

Democratisation of expertise represents an institutional effort to include actors previously excluded from consultation within the decision-making processes. It also focuses on participation in an institutional setting (e.g. science committees and *ad hoc* consultations), usually at the invitation of political authorities, and aims to foster 'transparency' and 'public information'. Finally, it recommends 'broadening expertise' to non-academic researchers and practitioners, 'establishing guidelines for the selection of the experts' appointed in science committees (Carolan 2008). Examples of the application of this approach can be found in a number of participatory technological assessments. At the invitation of the scientific institutions, these approaches aim to achieve a scientific consensus on technical and social questions by arguing between experts, counter-experts (Pellizzoni 2001, 2003; Abels 2007), civil society (Ferretti and Pavone 2009) and lay public perspectives. In these terms, participation focuses on bringing citizens closer to the decision-making by accounting for perspectives and expertise that would otherwise be ignored.

From a democratisation of research perspective, the potential for the lay public to contribute to scientific knowledge and scientific policy-making processes is recognised:

Most Committees value input from lay members; it is often their questions which open up issues which experts might not otherwise have explored. (UK Cabinet Office 1999)

The importance of users and stakeholders has also recently been acknowledged in contemporary innovation dynamics. In fact, user-centred innovation is a growing phenomenon that Von Hippel (2005) interpreted as a clear sign of a shift towards a democratisation of research innovation.

This growing trend towards democratisation of research and innovation represents the background that the report *Public Engagement in Science* (2008) suggested that growing hostility towards science and technology was due, not only to public failure to understand scientific risk objectively, but also to the narrow definition of problems and issues from within the scientific and policy-making processes. The increasing commercialisation of science, for instance, was accused of threatening scientific independence (Levidow 2002; Lotter 2008). The persistent attitude of treating the development and regulation of technology as merely 'scientific issues' and therefore

for scientists only, was also considered problematic (Felt and Wynne 2007).

A number of academic contributions have focused on the importance of public engagement with research by attempting to critically outline the complex dimensions of participative science (Nowotny 2003; Leshner 2003; Elam and Bertilsson 2003; Wilsdon and Willis 2004; Leach et al. 2005; MacNaghten et al. 2005; Wynne 2006). Other works focused on the ideal counterpart, scientific citizenship (Irwin 2001; Rose and Novas 2005; Rose 2007). In spite of the variety of different approaches to public participation and engagement in science, there seems to be an emerging consensus towards a shift from downstream assessment to upstream engagement and decision-making. Extensively elaborated by Wilsdon and Willis (2004), the upstream approach suggests that, in order to incorporate lay expertise, concerns and priorities, it is necessary to include public engagement at the early stage of the science policy decision-making process, ideally when the research agenda is being discussed and planned. This is preferable to engagement at the downstream level, where research has already been developed and where only its social, ethical and legal implications need to be assessed and discussed.

Upstream public engagement, however, poses new and equally challenging problems. First, little has been actually suggested or implemented to effectively involve civil society organisations (CSOs) and other stakeholders at the upstream level. Second, within the new vision of a neo-liberal state, which has been finely captured by the definition of a ‘competition state’, government action is no longer directed at the provision of welfare services nor confined to law, order or the protection of the market mechanisms. In fact, in the competition-state approach, the state actively fosters science and innovation dynamics by helping to develop emerging technologies and new market opportunities arising from the knowledge economy (Benner and Löfgren 2007). In other words, the state action is explicitly oriented towards both science and innovation stimulating economic growth and enhancing competitiveness. Under this assumption, however, even when all stakeholders are involved in upstream public engagement, not all positions and approaches stand on equal epistemic and economic ground. This is especially true for CSOs, where the main approach is oriented towards a science for the public good, rather than for economic growth and competition. In this context, sound scientific proposals which aim to investigate important scientific issues with social implications, but are unlikely to deliver patents or directly enhance economic growth, have less chance of gaining funding than equally sound proposals that are directly aimed at such goals. For this reason, it becomes crucial to ensure the active and effective participation of research stakeholders not only in the process of science policy-making, but also in the process of research evaluation.

3. Research evaluation: Towards stakeholder engagement

The use of bibliometrics as a tool for the democratisation of research evaluation is an argument frequently overlooked in favour of debating the benefits and drawbacks of peer review versus metrics. The following argument does not focus on pitching peer reviews against metrics but instead, draws on the advantages and disadvantages of both to demonstrate the value of including bibliometrics in research evaluation processes as part of a move towards the democratisation of research. Table 1 presents the main advantages and disadvantages of both bibliometrics and peer review as they are currently debated in the literature (Butler 2008; Taylor 2011; Archambault et al. 2006; Bornmann 2008; Moed 2007). The reader is encouraged to use this table to navigate the discussion below and to inform their own decision regarding the value of each tool in relation to the democratisation of research evaluation.

3.1. The public problems with peer review

Traditionally, research evaluation for the allocation of public funds has been limited to peer review, the validity of which is often called into question. While evaluation by peers ensures that scientific work is evaluated competently, it confines scientific evaluation to a minute, hyper-specialised committee of ‘experts’, operating behind closed doors. Any process whereby the recipients of public money are determined by a select group of peers, separate from the non-academic community, leads to questions about the validity of the resulting evaluations. It can, therefore, be unclear to disciplinary outsiders (including the public) why certain research projects are funded and others are not. Additionally, in times of limited resources and increasing societal demands, decisions from such a non-transparent process can be questioned by those who are unaware of the weighting of criteria or what aspects of research are valued.

One of the most important prerequisites for effective peer review is the motivation, competences and independence of the evaluators (Hemlin 1996; Kostoff 1997). At the same time, maintaining the quality of peer review is a time consuming and expensive process. Additionally, the risks of inadequate peer review, especially those that do not take into account growing public expectations from research, occur in countries with small academic communities and that are influenced by bias and misunderstandings, can be potentially devastating for the careers of individual researchers and the progress of science. In many grant, promotional and institutional ranking exercises, assessments are typically undertaken by a small number of assigned peers or by relatively small peer panels. The process is often conducted behind closed doors and judgments are based on a series of intellectual and

Table 1. Advantages and disadvantages of peer review and metrics

	Peer review	Bibliometrics
Is there an ability to consider content of research (including mitigating factors)?	Yes	No
Is there an ability to consider the potential future impact of researcher (ie. How research will perform in 3-4 years)	Yes	No
Does it consider up-to-date knowledge at the time of review?	Yes	No
Does it protect evaluations from an implicit bias	No	Yes
Does the halo effect exist?	Yes	No
Are the evaluation results accessible to non-academic stakeholders?	No	Yes
Are the evaluation results accessible to researchers in different fields?	No	Yes
Is the process time consuming?	Yes	No
Is the process expensive?	Yes	No
Does it produce replicable and reliable outcomes and results?	No	Yes
Is there international accessibility regardless of original language of publications?	No	Yes
Are all types of publications considered equally?	Yes	No
Is it applicable to all disciplines?	Yes	No
Does it consider the local-level relevance of the research?	Yes	No
Is it able to consider the societal impact of the research and/or researcher?	Yes	No
Is there protection from external social, economic and political pressure?	No	Yes
Does a conservative bias exist?	Yes	No

social processes that may be mediated by factors other than the quality, importance or impact of the research being evaluated (Langfelt 2001; Cicchetti 1991). For example, evaluators may be influenced by political and social pressures within the scientific community, such as possible repercussions from their decisions affecting their work (competition) and that of colleagues (Stehbens 1999). In addition, peer reviewers tend to evaluate work in terms of their own research interests and may not possess the granular knowledge required for expert analysis outside their immediate field, nor the broader knowledge needed for making 'big picture' judgments about the importance of research.

As such, the reliability of peer review is also questioned. As Langfeldt (2001) argues:

... while there is a certain set of criteria that reviewers pay attention to – more or less explicitly – these criteria are interpreted or operationalised differently by various reviewers.

Studies of peer review have shown that often its outcomes are not replicable (Cicchetti 1991). This is somewhat ironic, considering that the peer-review process was actually developed to increase the probability that other researchers could replicate reviewed research in order to give a measure of scientific quality control. Recent critics of peer review have also focused on the potential shortcomings of the reviewers, who, when subjected to strong social, economic or political pressures, may become:

... mean spirited, lead footed, capricious toadies and hacks that hide behind the cloak of anonymity. (Suls and Martin 2009)

The peer-review system, when used as the sole method to evaluate research, runs counter to the acceptable criteria of transparency and accountability apparent in many

government policies promoting the importance of the democratisation of research. Indeed, the current intense debate surrounding the future of adequate research evaluation balances the peer-review tradition with the desire of policy-makers (and ultimately the public) to ensure that publicly funded research yields the best scientific value and international relevance for the taxpayer's money (Taylor 2011). Consequently, many authors have lamented the existence of 'the relevance gap', where there is a disjunction between the research that society needs and the research being promoted as 'excellent' by peer review committees (Nightingale and Scott 2007). Frederiksen et al. (2003) stated that the public no longer trusts a narrow evaluation approach conducted in an isolated system with the one-dimensional and intra-scientific quality criteria that is peer review. From a knowledge production point of view, traditional peer review is better suited to the Mode 1 of research production. As research moves towards Mode 2, where research becomes reflexive, more complex and participatory research evaluation methods are required (Nowotny et al. 2001). These methods need to be socially reflexive in that they need to be accessible and understandable by a greater number of citizens outside the academic or disciplinary community. Additionally, as Mode 2 knowledge requires more socially relevant outcomes not sufficiently apprised by peer review alone, more complex methods accounting for new considerations of research excellence are required. Recent increases in the number of criteria to evaluate excellence in research, are seen as an attempt to create public trust in research (Nightingale and Scott 2007). Consequently, to increase public trust a subsequent increase in complexity is required. The main challenge for future evaluation systems, therefore, is to re-establish public trust in research while exploring new, trustworthy

forms of evaluation that are able to reach the various non-academic groups who participate in the discussions about scientific knowledge: scientists, policy-makers, the media, social movements, non-governmental organisations, business people, and lay citizens. In this way at least, and despite continual debate, bibliometrics can play a role.

3.2. A role for bibliometrics in democratic research evaluation

Bibliometrics is the study of the dissemination of scientific publications within the academic community and offers a powerful toolkit for studying the processes and achievements of research (Pritchard 1969). The observation that citations indicate use and therefore usefulness as well as impact is the basic argument for their use as an indicator of quality (Gläser and Laudel 2007). For research evaluation and the purposes of this paper, the attraction of bibliometrics lies not only in their transparency and ease of understanding, but also in their ability to translate information about research outcomes to non-academic stakeholders.

Using bibliometrics, non-academic stakeholders in the research process including the public, government decision-makers and other associated non-research personnel are able to deduce information regarding the perceived validity and impact of a piece of research, and research groups. Indeed, metrics and metric-informed rankings enjoy a high level of acceptance among stakeholders and wider public because of their simple and accessible information (Rauhvargers 2011). This, essentially, allows non-academic stakeholders to become active and equal participants in the research process in line with aspirations for the democratisation of expertise and of science. Further, the simplicity of bibliometrics allows for a greater understanding by these non-academic stakeholders about how research funding is distributed.

It must be acknowledged, however, that the use of bibliometrically informed rankings and scores has led to misunderstandings regarding the definition of research quality (Rauhvargers 2011). Currently, there are a number of university rankings providing different 'scores' and 'rankings', all claiming to represent 'prestige' and 'quality'. The oversupply of information can result in confusion for stakeholders, many of whom rely on this information as a guide to making decisions, for instance philanthropic donations or study choices. However, the difference between rankings that are informed metrically, and those that depend solely on the decisions of a specialised elite, is that the manipulation or bias of metric-informed rankings is easily identified externally. Indeed, formally incorporated metrics cannot be manipulated or 'conveniently misunderstood' by those who evaluate. Instead, if at all, manipulation must occur at the level of the metric creation and, is therefore, not only

traceable and identifiable by external stakeholders, but it can also be openly discussed in a debate.

In this respect, bibliometrics can compensate for one of the major problems associated with peer review, which is necessarily selective in the amount of research output that can be reviewed in the time span available. Bibliometric methods are capable of providing, albeit summarised, information about all research outputs for which citations are available (Abramo 2009). This openness of information not only decreases the possibility of conflicts of interest and corruption, but allows stakeholders to make decisions using indicators that incorporate aspects of research that are important for them, as taxpayers and consumers of information. In addition, when bibliometrics are incorporated into the research evaluation process, the extent that community members are capable of engaging with the decision-making process is no longer restricted to their understanding of the intricacies of the scientific field or membership of a minute 'expert group'.

Additional concerns about peer review further outline the importance of an official role for bibliometrics alongside peer review in research evaluation. In many peer-review evaluations, researchers are forced to make judgments about research that is beyond their area of expertise. This can increase the incidence of false positives, where the decision to award funding is made on the basis of an incorrect interpretation of research excellence. It is understandable that peer-review panels are time constrained. However, in such instances, not all research outputs can be evaluated in sufficient detail to ensure adequate evaluation (Taylor 2011). This effect is amplified in situations where the evaluators are assessing proposals outside of their direct discipline. In these cases, decisions are likely to be influenced by factors such as: knowledge of the author's previous publications, the status of the journal in which a publication appears, and the status of the author's department or university (Mingers and Xu 2010). In these cases, the 'halo effect' or an implicit bias (Taylor 2011; Reich et al. 2007) is used to evaluate instead. Using this logic, bibliometric variables already, in some way, informally influence the peer-review decision. However, the problem remains that while peer-review outcomes are influenced by metric variables, it is not clear which combination or weighting of such variables contribute to these decisions. If peer-review decisions are already informally using these objective measures to produce subjective decisions, then this could jeopardise any future formal role for bibliometrics in research evaluation. Indeed, it is precisely the casual and unstructured use of bibliometrics that negatively influences their reputation amongst researchers and, as such, their formal acceptance as tools for research evaluation. Bornmann (2008) states that metrics will only ever become a reliable research evaluation tool, if the resulting evaluations correlate with those made using peer review (Bornmann et al. 2008). However, a number of studies already confirm

peer-review decisions with a variety of commonly used bibliometric indicators (Taylor 2011; Bornmann et al. 2008; Derrick et al. 2011; Groot and Garcia-Valderrama 2006; Bornmann and Daniel 2006; Rinia et al. 1998). These studies demonstrate that peer-review decisions can be replicated or even predicted using bibliometrics and that this correlation may be stronger for some criteria and fields (Rinia et al. 1998). This result only further outlines how evaluators are already informally using metrics to inform their decisions. Bibliometrics can also provide information independent of the social prestige of authors or their affiliation with organisations with reputations for high quality research. By minimising this halo effect, bibliometrics can therefore eliminate a large amount of the bias associated with peer review (Moed 2007).

Governments and bibliometric researchers are likely to continue to persist in the development of reliable quantitative measures for future evaluation frameworks as their benefits seemingly outweigh their methodological flaws. The necessity for greater public accountability will demand that research provides publicly accessible and understandable benchmarks using bibliometrics, to explain how research funding is distributed. It is therefore likely that in future there will be an evaluation framework where all research participants, irrespective of their backgrounds, are able to make judgments about the validity, achievements and trustworthiness of research using the transparent and boundary-crossing characteristics of metrics. This would help to raise the profile of high quality research and researchers with the public and, at the same time, promote an open and accessible decision-making process for funding allocation.

The argument presented here does not overlook the inherent limitations (or modalities) of bibliometrics, a common concern of some bibliometric sociologists (Gläser and Laudel 2007). Instead it considers bibliometrics as an informative tool rather than adds to the debate about the current limitations of bibliometrics. The limitations of bibliometrics, however, can be separated into two categories: limitations that can be overcome as the field matures, and those that cannot. In the first category, these limitations depend on the maturity of the field and current capabilities of citation database providers. They include: the inadequacy of bibliometric databases to provide citations for fields such as the humanities (Pontille and Tornø 2010), and the failure of current indicators to reflect wider research impact (Taylor 2011). In both cases, as databases and considerations of research output become more comprehensive, and indicators more sophisticated, the impact of these limitations will diminish over time. Conversely, there are limitations associated with bibliometrics that are more difficult, if not impossible, to overcome. These include the use of a 'citation' as a proxy for 'research quality', and the assumption that publications reflect research productivity rather than just a small part of the research's overall value. Nevertheless, our

argument is not one that replaces peer review with bibliometrics, but that their limitations should not overshadow the implicit benefits of increased bibliometric use towards the reduction of some of the bias associated with peer review and a greater democratisation of research evaluation. Peer review at its best is an important component of the scientific process: but at its worst, it may turn out to be an obscure, unintelligible practice supporting self-serving groups of dominant academic scholars and institutions. However, the importance of the peer-review process for funding allocation being open and transparent is emphasised in the light of the ongoing quest for a more democratic and participative system of governance for science. Instead, we envision a new, more complex system that, through the successful integration of both evaluation tools, identifies research excellence and is accessible to a variety of new, previously excluded actors in the research evaluation process. Bibliometrics provides a common language to promote the openness of decision-making within the scientific community. Despite these advantages, formal government evaluation procedures and policies still rely on a narrow evaluation system that is, by nature, self-serving and runs contrary to the shift towards greater public participation in research. We outline three examples of this contradiction below.

4. Government frameworks incorporating metrics

Despite a number of government policies emphasising the importance of democratisation and the greater public engagement in science, their research evaluation frameworks fail to reflect the rhetoric of the increased accessibility, accountability and transparency of science. If research evaluation policies are to be effective, they must be consistent with other government policies that highlight the importance of transparent and open research through greater science communication. The consideration of whether or not to incorporate metrics into formal research evaluation procedures is constantly debated in the light of recent changes to national evaluation programs in Australia, the UK and Spain. We now review government policies that aim to promote the democratisation of science and the evaluation of research.

4.1 The United Kingdom

An interest in science communication and its relationship between science and the public gained momentum with the publication of the Royal Society's Bodmer Report (Bodmer 1985). The report stated that the reasons for promoting science communication were both individual and collective and that science would benefit from increased communication in the form of greater public favour and funding (Stein 2003). In the UK a White

Paper from the Committee to Review the Contribution of Scientists and Engineers to the Public Understanding of Science, Engineering and Technology, known as the Wolfendale Report, was submitted (Wolfendale 1993). The purpose of the committee was:

...to review the steps currently being taken to equip and encourage professional scientists, engineers and research students to contribute to improved public understanding of science, engineering and technology, and to suggest how these might be improved consistent with available funding.

It recommended building public understanding into research council grant procedures as well as other suggestions of how to increase public participation in research (Wolfendale 1993). It also commissioned a 'best practice' guide by the Office of Science and Technology. The committee concluded that:

In principle, all who receive grants from public funds should accept a responsibility to explain to the general public what the grant is enabling, or has enabled them, to do and why it is important and how it fits into the broader area of knowledge. (Wolfendale 1993)

Wolfendale's (1993) main recommendation was to build this theme of the greater democratisation of research into the research councils' grant procedures. The five UK research councils all had different types of public understanding of science strategies, including: training in communication skills for scientists, school liaison officers, PhD students and the compulsory reporting of science communication activities by grant holders. Despite these strategies, the report and the research councils failed to incorporate this greater democratisation of research rhetoric into their research evaluation guidelines.

However, the UK has taken a leading role in the development of research evaluation frameworks. The Research Assessment Exercise (RAE) has been in operation since 1992 and its replacement, the Research Excellence Framework (REF), is due to be implemented in 2012–3. The RAE's aim was to:

...produce ratings of research quality which will be used by higher education funding bodies in determining the main grant for research to the institutions they fund.

Several assessment exercises under the RAE model have been conducted since its inception and although the exact practices have changed over time as more feedback on the process was received, its basic framework has not.

The RAE was essentially, a peer-review system, assessing the 'quality' of research outputs using peer-review panels. Bibliometrics although not a formal constitute of the framework, no doubt informally contributes to the overall assessment of these outputs as a summarising and/or 'halo effect'. RAE's peer review panel basis involved more than 50,000 academic staff on up to 69 subject panels, each with 12 members (Graham 2008).

This arduous process was both time consuming, complex and expensive. In fact, the expense of the entire exercise was the catalyst for developing a cheaper, quicker and more efficient research framework. To achieve this efficiency, it was believed that bibliometrics should play a role in the new framework.

Initially, the REF,¹ unlike the RAE, introduced a system based largely on metrics, partially to reduce the cost and burden of the previous exercise. Initially the REF was to be heavily based on bibliometrics but a subsequent academic outcry replaced this emphasis with 'peer review by expert panels', similarly to the previous RAE. Bibliometrics were seen as incapable of capturing user value and impact and providing reliable assessments (Butler and McAllister 2009). The Higher Education Funding Council for England, however, admits that these expert panels will no doubt use bibliometric indicators to decide results informally, as was the case with the RAE but in the REF model, the individual sub-panels will be provided bibliometric data by the Higher Education Funding Council for England and then each panel is to decide how, if, and to what extent, these bibliometrics are to be used to determine the evaluation (Butler and McAllister 2011). No standards, guidelines or benchmarks will be imposed on how bibliometrics are to be used by the members of these panels. Nor is it clear what blend of variables was used to guide this evaluation.

Previous studies have investigated the relationship between RAE framework results and a number of bibliometric indicators (Abramo and D'Angelo 2011; Levitt and Thelwell 2011; Van Raan et al. 2007). This has also been done for fields that are not traditionally, well represented by conventional bibliometric indicators. Butler and McAllister (2009, 2011) state that, despite no single model being applicable for across all fields, bibliometric indicators still have the potential to be applied to formal evaluation frameworks (Graham 2008; Moosa 2011). Since these and other studies have found strong correlations between the RAE (now REF) results and a number of bibliometric indicators, this reveals several policy implications for how to evaluate research transparently. First, as the correlation exists, this suggests that, to some extent, bibliometrics are being already used by peer-review evaluation committees when they make their decisions. Further, this continued informal use of metrics implies that bibliometric aspects are already considered, in part, sufficient to measure a modern definition of 'research excellence'. In fact, this informal use by evaluation panel members suggests that academics, at least, value the characteristics of research excellence that can be reflected by bibliometric indicators. The implicit bias already in practice by evaluation panels threatens the reliability and validity of these evaluations and more importantly, impairs the transparency of the evaluation outcome. By relying too heavily on a system that combines peer review with haphazardly applied bibliometrics, it becomes increasingly difficult to

discern how a decision was made and why. With neither an expert background nor a transparent system with which to understand these outcomes, it is even more difficult for non-academic stakeholders to understand why and where their taxpayer money is being distributed.

With a sufficiently advanced bibliometric methodology it is possible to construct an appropriate, coherent set of indicators for assessing research performance in a way that is: acceptable for the institutions and research groups that are to be evaluated; comprehensive across at least the science-based disciplines; and robust and reliable for specific research themes, fields and entire disciplines (Van Raan et al. 2007). The UK has a reputation for being innovative in the creation of research and innovation policies, especially in relation to public engagement and research evaluation. As bibliometrics become more sophisticated and each limitation is overcome with appropriate application and understanding, it remains to be seen whether the UK's commitment to policies in the democratisation of science will be mirrored in future research evaluation policies and frameworks.

4.2 Australia

Recent changes to the Australian research evaluation framework, the Excellence in Research for Australia (ERA), have downplayed the role bibliometrics plays in the assessment of research organisations.² Arguably, changes in the framework over time have threatened to challenge the democratisation of science, as the framework continually replaces the role of bibliometrics with a peer-review basis. Like the UK, Australia has proposed two research frameworks and has policies similar to those in the UK, which aim to increase the level of public participation in research. However, unlike the UK, the first framework supported a large role for metrics. This framework was then superseded by another that emphasised the role of panel-based peer review.

In Australia, government policy has emphasised a level of public accountability for scientists. The first official federal government policy for scientific research was only first established in 1999 with Backing Australia's Ability Policy (Department of Innovation, Science and Research 2001). This policy emphasised scientists' 'duty' to communicate their research if they were in receipt of public funds, a view that echoed the Royal Society Report (Bodmer 1985) in the UK. This commitment to aspects of the greater democratisation of research has been echoed in subsequent government science policies and has manifested itself in activities such as: National Science Week, the development of national research priorities, and the inclusion of questions regarding 'community outreach' activities of Australian Research Council (ARC) and National Health and Medical Research Council funding applications. However, as demonstrated above, a commitment to the greater democratisation of research through

policies about stakeholder engagement is not necessarily reflected in parallel research policies about research evaluation.

With regards to developing official research evaluation frameworks, Australia has mainly taken its lead from the UK. The ill-fated Research Quality Framework (RQF) heralded the first government commitment to the greater evaluation of research within publicly funded universities (Derrick 2011). The RQF was characterised by greater emphasis on the use of bibliometrics to guide evaluations. However, a large backlash from the academic community combined with a change in government, led to the scrapping of the RQF in favour of the ERA framework. The ERA favoured a more mixed methods approach combining peer review with metric calculations. It is the provision of the ARC Ranked Journal List, however, that will be the main focus of this discussion.

The ARC Ranked Journal List was part of the first round of the ERA in 2010. The ERA ranking included sorting 20,172 journals into four quality categories: A*, A, B, and C in such a way that A* represents the top 5% of journals, A the next 15%, B the next 30%, and C the remaining (lower) 50% of journals (Graham 2008). The ARC journal ranking list was developed in four stages that included public consultation and review by field 'experts'. This approach uses a survey- or perception-based method which presents subjective decisions objectively. In this way, the 'ranking' effectively relieves the peer-review process of the issue of transparency and minimises the effect of bias, although recent research has argued that more quantitative measures like rankings based on citations, are more open to bias than this survey-based approach (Moosa 2011). The result was a simply constructed list providing researchers and the public with a guideline on how to determine excellence and visibility in publishing from research and researchers. Until recently, the ERA process relied heavily on this journal ranking scheme to evaluate publicly funded research in universities. However, despite being implemented in the 2010 ERA, subsequent academic retaliation concerned about using journal ranking to evaluate research output, resulted in the scrapping of the list for the second ERA in 2012. The decision to scrap the list was, according to the Minister for Innovation, Industry, Science and Research, that:

... there is clear and consistent evidence that the rankings were deployed inappropriately within some quarters of the sector, in ways that could produce harmful outcomes. (Carr 2007)

These harmful outcomes included: potential bias resulting from the ranking of journals based on the responses of a potentially biased sample of academics; construction of the list was time consuming without a tangible benefit; they resulted in a 'race to the bottom' in the sense that scholars were motivated to publish in journals that were ranked A* but with little international recognition; an adverse effect

on the research dealing with national issues; and the encouragement of ‘unproductive rent-seeking activities’ as editors lobbied for high rankings for their journals and authors did the same for the journals in which they published. The Minister further elaborated that the ‘ERA could work perfectly well without the rankings’ and that:

...their [rankings]’ existence was focusing ill-informed undesirable behaviour in the management of research. (Carr 2007)

For ERA 2012, the ranking list was replaced with another list which included the journals to be considered but omitted the A*-C ratings. How exactly this list will be used to assist the ERA research evaluation committees is not yet clear. Instead, as with the UK REF process, the peer-review committees will employ undefined, scattered criteria to make a judgment about the quality of publications and research. In addition, while the description of the ERA objectives refer to the use of ‘indicators’ and ‘citations’, it is not clear how or to what extent these are used as a criteria in the evaluation process.

The solution may not be abolition, but rather refinement. The strength of the journal rating scheme was in its simplicity, level of peer-involvement and commitment and reflection of other government policies urging the greater democratisation of research and researchers. With bibliometric research increasingly producing new indicators and contributing to the research evaluation debate both within and beyond its traditional field of research (Jonkers and Derrick 2012), the potential for research evaluation frameworks and journal rankings that are both academically sophisticated and publicly accessible, is increasingly likely. Their formal integration is preferable over the current model where bibliometrics, or at least the perceived reputation of a journal, are applied haphazardly without guideline or benchmarks to ensure the transparency, validity and accountability of evaluation outcomes.

In addition, despite there being a current connection between funding and ERA outcomes, future plans indicate a larger role for ERA results in determining the allocation of taxpayer-funded research dollars to award research activities in the country’s universities (Hemlin 1996). With public funding being determined by these outcomes, it is more important than ever to ensure an evaluation scheme that is transparent and consistent with all government policies regarding the future of research and the engagement of non-academic stakeholders.

4.3 Spain

Research evaluation policy in Spain differs from that in both Australia and the UK as the importance of the ‘democratisation of science’ and the transparency of research has not been specifically stated in Spanish science and research policy (Ferretti and Pavone 2009; Pavone and Arias 2011). Instead, the main concern has been the fair

distribution of funding and the internationalisation of research (Sanz-Menéndez 1995; Cruz-Castro and Sanz-Menéndez 2007). In particular, the accountability of researchers is used to produce internationally recognised research. Accountability is an essential component of the democratisation of research and although public engagement has not been explicitly recognised in Spanish research policy, it is becoming increasingly implicit in research funding allocation and the evaluation of researchers. As such it represents the Spanish government’s approach to an increased democratisation of research.

This difference between Spanish and the UK or Australian policies is due to how the relationship between government, civil society and science has evolved over the last 20 years. The changes in research policy and an increased investment in R&D activities since the demise of Franco’s regime have not only changed the way that Spanish researchers operate but have also raised the question of how to evaluate these activities in the light of movement towards a greater accountability for Spanish research and researchers.

Since 1990, the Spanish government has introduced two major research evaluation policies. The first is conducted by the National Agency for the Evaluation and Foresight (in translation: Agencia Nacional Evaluación y Prospectiva (ANEP)).³ The agency’s main task is to review research proposals and assign public funding. Evaluations employ an extensive network of peers in committees to review the proposals. However, the role of the ‘peer committees’ is restricted to conducting the evaluation and is not directly linked to the final decision. This final decision is made by another, separate, panel also established by ANEP. Technically, ANEP may conduct re-evaluations of their decisions if requested, but this rarely occurs because the main function of ANEP is to distribute funds fairly, based on the proposal’s merit at the time of application, rather than to evaluate research outcomes of past proposals that were successful (Sanz-Menéndez 1995). During the period 1985–92, when the economy was expanding, the accountability of researchers was of little concern to ANEP. However, during times of economic recession (1992–7, and at present) economic rationalism of taxpayer-funded research money is prioritised. Despite this, at times when research output is remarkably good (as measured by Spain’s visibility in the Science Citation Index databases), little is done to emphasise the importance of evaluating the research outcomes of past projects (Jiménez-Contreras et al. 2003).

The National Commission for the Evaluation of Research Activity (in translation: Comisión Nacional Evaluación Actividad Investigadora (CNEAI))⁴ was created in 1990 as the second research evaluation body. Its purpose was to formally incentivise the improvement or maintenance of research productivity for researchers who had been awarded the security of ‘civil servant’

status. Instead of focusing on research groups and/or institutional outputs as in the UK and Australia, Spanish policy-makers opted to evaluate the research trajectory of individual researchers. The CNEAI is a voluntary process for researchers and, if successful, researchers are rewarded by a small salary increase (Osuna et al. 2011). CNEAI evaluations are carried out by peer-review committees, who evaluate the researcher's five best publications (items) from the last six years. In contrast to evaluation systems from the UK and Australia, however, bibliometrics play a large, more formal (structured) role in these decisions. This structured approach involves CNEAI evaluations adopting a ranking system that rewards researchers for articles published in the 'first quartile' of ISI journals in each sub-field within the Science Citation Index or the Social Sciences Citation Index (Jost et al. 2009). Obtaining visibility in these journals is considered to be a measure of international recognition, and thus of research excellence, by the CNEAI.

The introduction of CNEAI evaluations was reportedly responsible for the increase in the number of Spanish papers and their citations in the Science Citation Index since the 1990s (Jiménez-Contreras et al. 2003). This, however, has been challenged by reports that the CNEAI was only effective in its initial years and has only consolidated an already existing trend by researchers to publish in international, high-impact journals. This explanation is supported by a strong correlation between the number of publications and rising number of research staff (Osuna et al. 2011). Nonetheless, the CNEAI system has been successful in fostering an increased internationalisation of Spanish research; the importance of high quality publications; and an appreciation for the importance of research evaluation. By utilising bibliometrically informed evaluations, the CNEAI evaluation committees are provided with information about the candidate independent of more subjective judgments about their academic reputation (the halo effect). In addition, the utilisation of bibliometrics is particularly advantageous for countries like Spain, where proficiency in the English language is essential in order to publish in high-impact, internationally focused journals but is not possessed by all researchers, some of whom may be more nationally focused. In these cases, bibliometrics has the ability to provide information to members of a CNEAI peer-review committee where language differences may present an obstacle to the full appreciation of the quality of the candidate's publications and research.

However, despite the utilisation of bibliometrics by CNEAI committees, the role of traditional, subjective peer review still dominates the decision-making process. Therefore, this subjective, unstructured and non-transparent aspect of the decision-making process fails to guarantee the transparency of research evaluations and the accountability of evaluators emphasised by the policies that established the two evaluation bodies described

above. Additionally, despite representing an important step towards incorporating an advisable combination of peer review with bibliometrics, it does not constitute an explicit step forward towards the greater democratisation of science or of research evaluation.

Unlike the UK and Australia, Spain's commitment to the application of bibliometrics for research evaluation occurs independently of any explicit parallel government policy highlighting the importance of increasing the democratisation of research. Instead, the research policies focus on research accountability as one important aspect of the democratisation of research. Even though not explicitly mentioned in Spanish government policies, the concept of democratisation in the receipt of public funds is now gaining momentum, especially among the youngest members of the research community. In the absence of an explicit policy, however, Spain may still require further assistance to incorporate the concept of democratisation into formal research evaluations.

5. Conclusions

This paper has outlined the current contradictions between the efforts citing the benefits of the greater democratisation of science and the emerging debate surrounding the structure of research evaluation systems for the distribution of taxpayer-funded research grants. This was the case for the three nations discussed in Section 6, where policies emphasising the importance of accountability, openness and other factors associated with the greater democratisation of science are not reflected in parallel policies and practices regarding research evaluation.

The growing popularity of metrics appears to be never-ending, indeed the recent growth in the bibliometric literature (Jonkers and Derrick 2012) is likely to ensure that the politics surrounding the use of metrics will also remain a continual, almost chronic, concern endemic to the research evaluation debate. Their advantages include: providing information that is open, inexpensive and that transcends boundaries but these attributes have overshadowed their potential for fostering the greater democratisation of research evaluation. Instead a shift away from the habitual discussion of if bibliometrics should be used towards a more constructive dialogue about how they can be used is required. This shift in rhetoric is essential. Despite the level of bibliometric consideration, the problems associated with peer review in research evaluation will not disappear. Further, the prospect of ever overcoming these problems of peer review is unlikely in the absence of a constructive dialogue about the inclusion of alternative methods of research evaluation. Bibliometrics, at least, deserves this consideration.

In addition, the opportunity to increase participation of non-academic stakeholders, as in line with the ideals of the democratisation of science, is unlikely if the research

evaluation system devalues non-academic stakeholder opinions based solely on their segregation from this elite group of scientific members. The simple information portrayed by correctly calculated and applied bibliometric indicators has the potential to engage a larger group of stakeholders than previous evaluation systems could. These stakeholders would include researchers from other related fields, policy-makers, politicians and the public. These advantages of bibliometric use should not be discounted in favour of advocating against their use at all. By doing this, research contradicts its own rhetoric about the benefits of a wider, more publically engaged science by defending tradition that moves research further away from this goal. The inclusion of bibliometrics in research evaluation systems would facilitate the engagement of a wider range of stakeholders towards a more ideal model of 'stakeholder evaluation'. Our argument is not that peer review be replaced with bibliometrics, but rather that the ability of each method to compensate for the failings of the other promotes both the scientific and social advantages of a combined model. The operationalisation of this model is a topic that deserves more, widespread debate. Although problems can be envisioned with this combined model such as: what happens when peer review and bibliometric outcomes differ; and there could be an inadequate understanding of the modalities of bibliometrics by the evaluators. These concerns are worthy of discussion and can be addressed by a clear and balanced debate and a well-defined, structured role for both approaches to evaluation. This paper overwhelmingly supports the necessity of this dialogue.

With the discussion on research evaluation dominated by arguments about bibliometric inadequacy, government policy-makers can be partly forgiven for failing to synthesise policies considering the greater democratisation of science with research evaluation systems. Nonetheless, future policy discussions should not overlook the advantage of bibliometrics or the importance of greater democratisation of research when creating a more holistic, reflexive research evaluation policy.

Funding

This work was supported by the Spanish Ministry of Economy and Competitiveness (CSO2011-26019).

Acknowledgements

The authors would like to thank Emeritus Professor Chris Bryant from the Australian National University for his valuable insights during the production of this paper.

Notes

1. See <www.ref.ac.uk> accessed 13 March 2013.

2. See <www.arc.gov.au/era> accessed 13 March 2013.
3. See <http://www.idi.mineco.gob.es/portal/site/MICINN/menuitem.29451c2ac1391f1febebed1001432ea0/?vgnnextoid=3cb39bc1fccf4210VgnVCM1000001d04140aRCRD&lang_chosen=en> accessed 20 July 2012
4. See <http://www.educacion.gob.es/horizontales/ministerio/organigrama/ministro/relacionadas-administrativamente/cneai.html> accessed 10 November 2012.

References

- Abels, G. (2007) 'Citizen involvement in public policy-making: Does it improve democratic legitimacy and accountability? The case of pTA', *Interdisciplinary Information Sciences*, 13: 103–16.
- Abramo, G. and D'Angelo, C. A. (2011) 'National research assessment exercises: A comparison of peer review and bibliometrics rankings', *Scientometrics*, 89: 929–41.
- , — and Caprasecca, A. (2009) 'Allocative efficiency in public research funding: Can bibliometrics help?', *Research Policy*, 38: 206–15.
- Archambault, E., Vignola-Gagne, E., Cote, G., Lariviere, V. and Gingras, Y. (2006) 'Benchmarking scientific output in the social sciences and humanities: The limits of existing databases', *Scientometrics*, 68: 329–42.
- Armstrong, K. A. (2002) 'Rediscovering civil society: The European Union and the White Paper on Governance', *European Law Journal*, 8: 102–32.
- Benner, M. and Löfgren, H. (2007) 'The bio-economy and the competition state: Transcending the dichotomy between coordinated and liberal market economies', *New Political Science*, 29: 77–95.
- Bodmer, W. (1985) *The Public Understanding of Science*. London: Royal Society.
- Bornmann, L. (2008) 'An analysis of the peer review process from the perspective of sociology of science theories', *Human Architecture: Journal of the Sociology of Self Knowledge*, VI/2: 23–38.
- and Daniel, H. D. (2006) 'Selecting scientific excellence through committee peer review: A citation analysis of publications previously published to approval or rejection of post-doctoral research fellowship applicants', *Scientometrics*, 68: 427–40.
- , Wallon, G. and Ledin, A. (2008) 'Is the h index related to (standard) bibliometric measures and to the assessments by peers? An investigation of the h index by using molecular life sciences data', *Research Evaluation*, 17: 149–56.
- Butler, L. (2008) 'Using a balanced approach to bibliometrics: quantitative performance measures in the Australian Research Quality Framework', *Ethics in Science and Environmental Politics*, 8: 83–92.
- and McAllister, I. (2009) 'Metrics or peer review? Evaluating the 2001 UK Research Assessment Exercise in political science', *Political Studies Review*, 7: 3–17.
- and — (2011) 'Evaluating university research performance using metrics', *European Political Science*, 10: 44–58.
- Carolan, M. S. (2008) 'Democratizing knowledge', *Science, Technology & Human Values*, 33: 508–28.
- Carr, K. (21 December 2007) 'Cancellation of research quality framework implementation'. Media release from Senator the Hon Kim Carr, Minister for Innovation, Industry, Science and Research. Canberra, Australia: ACT, <www.arc.gov.au/media/media-releases.htm> accessed 13 March 2013.

- Cicchetti, D. V. (1991) 'The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation', *Behavioral and Brain Sciences*, 14: 119–86.
- Cruz-Castro, L. and Sanz-Menéndez, L. (2007) 'Research evaluation in transition'. In: Whitley, R. and Gläser, J. (eds) *The Changing Governance of the Sciences*, pp. 205–23. Dordrecht, the Netherlands: Springer.
- Derrick, G. E., Haynes, A. S., Chapman, S. and Hall, W. D. (2011) 'The association between four citation metrics and peer rankings of research influence of Australian researchers in six fields of public health', *Plos One*, 6/4: e18521.
- Derrick, G. (2011) 'Trends and recent changes in the funding arrangements of Australian university research', *Informe CYD 2010, La contribución de las universidades españolas al desarrollo*, pp. 133–6. Madrid, Spain: Fundacion Conocimiento y Desarrollo.
- Department of Innovation, Science and Research. (2001) 'Backing Australia's ability'. Canberra, Australia: ACT.
- Elam, M. and Bertilsson, M. (2003) 'Consuming, engaging and confronting science', *European Journal of Social Theory*, 6: 233–51.
- Eriksen, E. O., Joerges, C. and Neyer, J. (2003) *European Governance, Deliberation and the Quest for Democratisation*. Oslo: ARENA.
- European Commission. (2008) 'Public Engagement in Science – Report of the Science in Society Session', European Commission: Brussels, <http://ec.europa.eu/research/science-society/document_library/pdf_06/public-engagement-081002_en.pdf> accessed 14 March 2013.
- Felt, U. and Wynne, B. (2007) 'Taking European knowledge society seriously' EUR 22:700. Luxembourg: Directorate-General for Research, European commission.
- , Fochler, M. and Winkler, P. (2010) 'Coming to terms with biomedical technologies in different technopolitical cultures: A comparative analysis of focus groups on organ transplantation and genetic testing in Austria, France, and the Netherlands', *Science, Technology & Human Values*, 35: 525–53.
- Ferretti, M. P. and Pavone, V. (2009) 'What do civil society organisations expect from participation in science? Lessons from Germany and Spain on the issue of GMOs', *Science and Public Policy*, 36: 287–99.
- Frederiksen, L. F., Hansson, F. and Wenneberg, S. B. (2003) 'The Agora and the role of research evaluation', *Evaluation*, 9: 149–72.
- Funtowicz, S. O. and Ravetz, J. R. (1993) 'Science for the post-normal age', *Futures*, 25: 739–55.
- Gläser, J. and Laudel, G. (2007) 'The social construction of bibliometric evaluations'. In: Whitley, R. and Gläser, J. (eds) *The Changing Governance of the Sciences*, pp. 101–23. Dordrecht, the Netherlands: Springer.
- Graham, L. J. (2008) 'Rank and file: Assessing research quality in Australia', *Educational Philosophy and Theory*, 40: 811–5.
- Groot, T. and Garcia-Valderrama, T. (2006) 'Research quality and efficiency: An analysis of assessments and management issues in Dutch economics and business research programs', *Research Policy*, 35: 1362–76.
- Hemlin, S. (1996) 'Research on research evaluation', *Social Epistemology: A Journal of Knowledge, Culture and Policy*, 10: 209–50.
- Irwin, A. (2001) *Sociology and the Environment: A Critical Introduction to Society, Nature, and Knowledge*. Cambridge, UK: Polity.
- Jasanoff, S. (2004) *States of Knowledge: The Co-Production of Science and Social Order*, London, UK: Routledge, Taylor & Francis Group.
- Jiménez-Contreras, E., de Moya Anegón, F. and López-Cózar, E. D. (2003) 'The evolution of research activity in Spain: The impact of the National Commission for the Evaluation of Research Activity (CNEAI)', *Research Policy*, 32: 123–42.
- Jonkers, K. and Derrick, G. E. (2012) 'The bibliometric bandwagon: Characteristics of bibliometric articles outside the field literature', *Journal of the American Society for Information Science and Technology*, 63: 829–36.
- Jost, J. T., Rudman, L. A., Blair, I. V., Carney, D. R. et al. (2009) 'The existence of implicit bias is beyond reasonable doubt: A refutation of ideological and methodological objections and executive summary of ten studies that no manager should ignore'. In: Brief, A. P. and Staw, B. M. (eds) *Research in Organizational Behavior*, 29, pp. 39–69. New York: Elsevier.
- Kerr, A., Cunningham-Burley, S. and Amos, A. (1997) 'The new genetics: Professionals' discursive boundaries', *Sociological Review*, 45: 279–303.
- Kostoff, R. N. (1997) 'The principles and practices of peer review', *Science and Engineering Ethics*, 3: 19–34.
- Langfeldt, L. (2001) 'The decision-making constraints and processes of grant peer review, and their effects on the review outcome', *Social Studies of Science*, 31: 820–41.
- Leach, M., Scoones, I. and Wynne, B. (2005) 'Introduction: science, citizenship and globalization', *Science and Citizens: Globalization and the Challenge of Engagement*, pp. 3–14. London: Zed.
- Lengwiler, M. (2008) 'Participatory approaches in science and technology', *Science, Technology & Human Values*, 33: 186–200.
- Leshner, A. I. (2003) 'Public engagement with science', *Science*, 299/5609: 977.
- Levidow, L. (2002) 'Marketizing higher education: Neoliberal strategies and counter-strategies', *Education and Social Justice*, 3: 12–24.
- Levitt, J. M. and Thelwell, M. (2011) 'A combined bibliometric indicator to predict article impact', *Information Processing & Management*, 47: 300–8.
- Liberatore, A. and Funtowicz, S. (2003) 'Democratising expertise, expertising democracy: What does this mean, and why bother?' *Science and Public Policy*, 30: 146–50.
- Lotter, D. (2008) 'The genetic engineering of food and the failure of science, part 2: Academic capitalism and the loss of scientific integrity', *International Journal of the Sociology of Agriculture and Food*, 16: 50–68.
- MacNaghten, P., Kearnes, M. B. and Wynne, B. (2005) 'Nanotechnology, governance, and public deliberation: What role for the social sciences?' *Science Communication*, 27: 268–291.
- Mingers, J. and Xu, F. (2010) 'The drivers of citations in management science journals', *European Journal of Operational Research*, 205: 422–30.
- Moed, H. F. (2007) 'The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review', *Science and Public Policy*, 34: 575–83.
- Moosa, I. (2011) 'The demise of the ARC journal ranking schemes: As ex post analysis of the accounting and finance journals', *Accounting and Finance*, 51: 809–36.
- Nightingale, P. and Scott, A. (2007) 'Peer review and the relevance gap: Ten suggestions for policy-makers', *Science and Public Policy*, 34: 543–53.
- Nowotny, H. (2003) 'Democratising expertise and socially robust knowledge', *Science and Public Policy*, 30: 151–6.
- , Scott, P. and Gibbons, M. (2001) *Re-thinking Science: Knowledge and the Public in an Age of Uncertainty*. Cambridge, UK: Polity.

- Osuna, C., Cruz-Castro, L. and Sanz-Menendez, L. (2011) 'Overturning some assumptions about the effects of evaluation systems on publication performance', *Scientometrics*, 86: 575–92.
- Pavone, V. and Arias, F. (2011) 'Beyond the geneticization thesis: The political economy of PGD/PGS in Spain', *Science, Technology & Human Values*, 37: 235–61.
- Pellizzoni, L. (2001) 'The myth of the best argument: Power, deliberation and reason', *British Journal of Sociology*, 52: 59–86.
- . (2003) 'Knowledge, uncertainty and the transformation of the public sphere', *European Journal of Social Theory*, 6: 327–55.
- Pontille, D. and Torny, D. (2010) 'The controversial policies of journal ratings: Evaluating social sciences and humanities', *Research Evaluation*, 19: 347–60.
- Pritchard, A. (1969) 'Statistical bibliography or bibliometrics?' *Journal of Documentation*, 24: 348–9.
- Rauhvargers, A. (2011) 'Global university rankings and their impact', *EUA Report on Rankings 2011*, p. 12. Brussels: European University Association.
- Reich, D. A., Green, M. C., Brock, T. C. and Tetlock, P. E. (2007) 'Biases in research evaluation: Inflated assessment, oversight or error-type weighting?' *Journal of Experimental Psychology*, 43: 633–40.
- Rinia, E. J., van Leeuwen, T. N., van Vuren, H. G. and van Raan, A. F. J. (1998) 'Comparative analysis of a set of bibliometric indicators and central peer review criteria: Evaluation of condensed matter physics in the Netherlands', *Research Policy*, 27: 95–107.
- Rose, N. (2007) *The Politics of Life Itself. Biomedicine, Power and Subjectivity in the 21st Century*. Princeton, NJ: Princeton University Press.
- and Novas, C. (2005) 'Biological citizenship'. In: Ong, A. and Collier, S. J. (eds) *Global Assemblages*, pp. 439–63. Malden, MA: Blackwell.
- Rowe, G. and Frewer, L. J. (2005) 'A typology of public engagement mechanisms', *Science, Technology & Human Values*, 30: 251–90.
- , Horlick-Jones, T., Walls, J., Poortinga, W. and Pidgeon, N. F. (2008) 'Analysis of a normative framework for evaluating public engagement exercises: Reliability, validity and limitations', *Public Understanding of Science*, 17: 419–41.
- Sanz-Menéndez, L. (1995) 'Research actors and the state: Research evaluation and evaluation of science and technology policies in Spain', *Research Evaluation*, 5: 79–88.
- Saurugger, S. (2010) 'The social construction of the participatory turn: The emergence of a norm in the European Union', *European Journal of Political Research*, 49: 471–95.
- Stehbens, W. E. (1999) 'Basic philosophy and concerns underlying scientific peer review', *Medical Hypotheses*, 52: 31–6.
- Stein, J. A. (2003) 'Public understanding of science and the policy context in the United Kingdom'. In: Felt, U. (ed.) *OPUS 5th Framework Programme - Raising Public Awareness of science and Technology in Europe*, Chapter 12. Vienna: University of Vienna.
- Suls, J. and Martin, R. (2009) 'The air we breathe: A critical look at practices and alternatives in the peer-review process', *Perspectives on Psychological Science*, 4: 40–50.
- Taylor, J. (2011) 'The assessment of research quality in UK universities: Peer review or metrics?' *British Journal of Management*, 22: 202–17.
- Todt, O. and Luján, J. L. (2000) 'Spain: commercialization drives public debate and precaution', *Journal of Risk Research*, 3: 237–45.
- van Raan, A., Moed, H. and van Leeuwen, T. (2007) 'Scoping study on the use of bibliometric analysis to measure the quality of research in UK higher education institutions'. Report to HEFCE. Leiden, the Netherlands: Leiden University.
- UK Cabinet Office. (1999) 'Modernising government', *London: Cabinet Office*.
- Von Hippel, E. (2005) *Democratizing Innovation*. Cambridge, MA: MIT Press.
- Weingart, P. (1999) 'Scientific expertise and political accountability: Paradoxes of science in politics', *Science and Public Policy*, 26: 151–61.
- Wilsdon, J. and Willis, R. (2004) *See-through Science: Why Public Engagement needs to Move Upstream*. London: Demos.
- Wolfendale, A. W. (1993) 'Realising our potential: A strategy for science, engineering and technology' Report by Committee to review the contribution of scientists and engineers to the public understanding of science. London: The Stationery Office London, <<http://www.official-documents.gov.uk/document/cm22/2250/2250.asp>> accessed 13 March 2013.
- Wynne, B. (2006) 'Public engagement as a means of restoring public trust in science: Hitting the notes, but missing the music?', *Public Health Genomics*, 9: 211–20.