SPRU
Science and Technology Policy Research

SEWPS

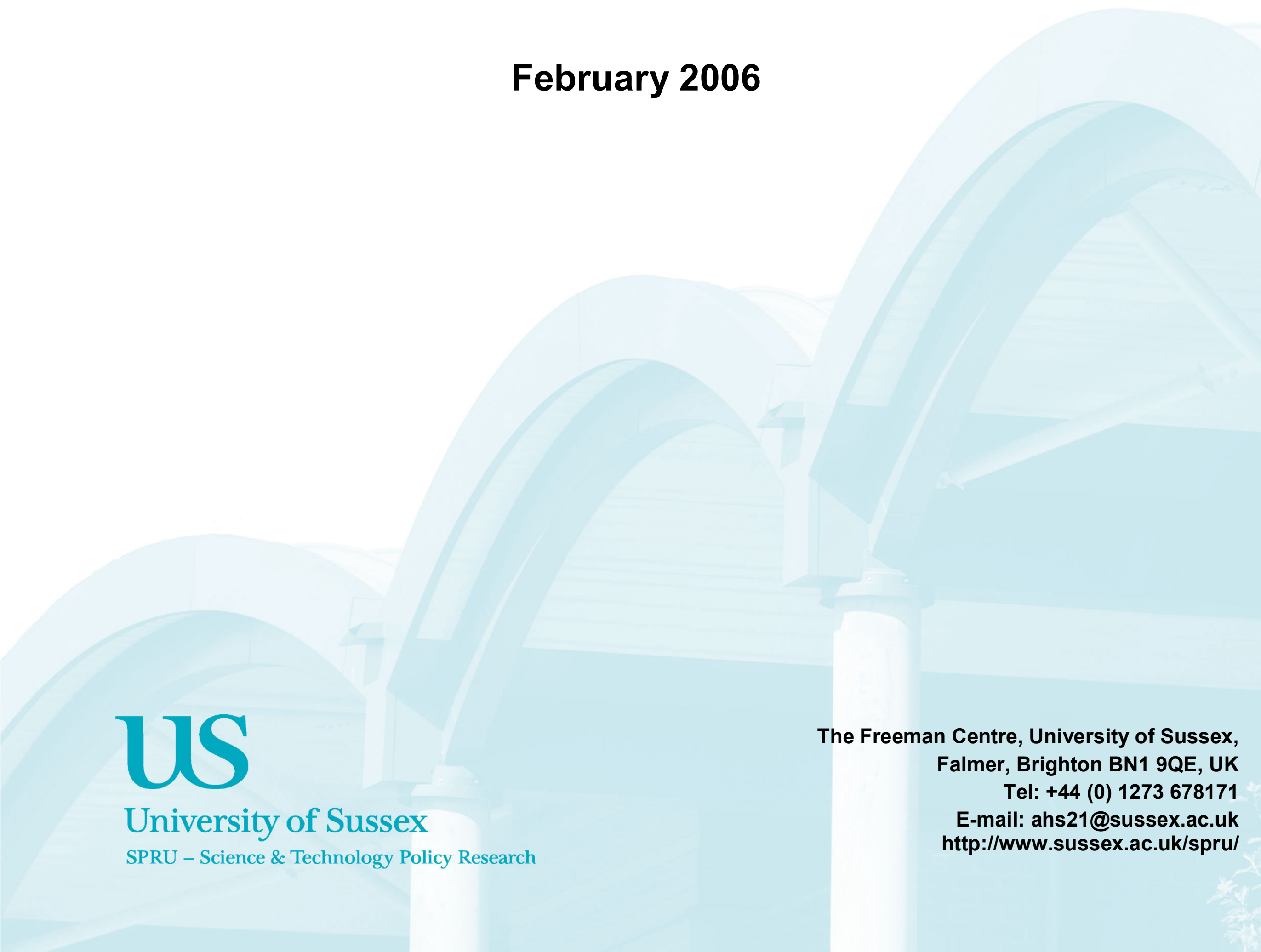
SPRU Electronic Working Paper Series

Paper No. 145

Peer Review and the Relevance of Science

Alister Scott
(SPRU)

February 2006



University of Sussex
SPRU – Science & Technology Policy Research

The Freeman Centre, University of Sussex,
Falmer, Brighton BN1 9QE, UK
Tel: +44 (0) 1273 678171
E-mail: ahs21@sussex.ac.uk
http://www.sussex.ac.uk/spru/

PEER REVIEW AND THE RELEVANCE OF SCIENCE

Alister Scott, SPRU, University of Sussex
Dr Alister Scott is Research Fellow at SPRU and partner at TheKnowledgeBridge Ltd.Freeman Centre, University of Sussex, Falmer, Brighton, BN1 9QE
Tel: (direct) 01273 678171, Fax: 01273 685865; E-mail: ahs21@sussex.ac.uk

Abstract

Recent science-policy debates have emphasised a growing role for science in helping to address some of society's most pressing challenges such as global environmental change, caring for the needs of ageing populations, and competitiveness in a global age. Other 'relevance' pressures include drives for public accountability, pressure for the 'democratisation' of science and demands from industry for usable knowledge.

Underlying the question of the social relevance of science is the matter of decision-making and quality control in science, usually via the peer-review process. Peer review plays a central role in many of the key moments in science. It is the main form of decision-making around grant selection, academic publishing and the promotion of individual scientists within universities and research institutions. It also underpins methods used to evaluate scientific institutions.

Yet peer review as currently practised can be narrowly scientific, to the exclusion of other pressing quality criteria relating to social relevance. It is often also controlled and practised by scientists to the exclusion of wider groups that might bring valuable perspectives.

This article sets out to examine peer review through the lens of social relevance. It challenges peer review as currently practised and makes some suggestions for ways forward.

Keywords: science policy, relevance of science, social relevance, peer review, quality control.

Abstract: 199 words

TABLE OF CONTENTS

PEER REVIEW AND THE RELEVANCE OF SCIENCE

Introduction

> 'Peer review is not only a routine component of the scientific role, but it is also fundamental to the institution of science, defended as symbol and guarantor of the autonomy of science. Thus peer review is built so deeply into the brickwork of science that many refuse to examine and improve it, fearing that any significant change would weaken the entire edifice. In some minds, to question peer review is to question science itself…' (Chubin and Hackett 1990: 2).

Since this thought from Chubin and Hackett – and possibly for the reasons they identify – we do not seem to have made much progress in studying the practice of peer review. My investigations of the procurement of 'relevant' science have led me inexorably to the topic of peer review, and to some critical perspectives on how peer review, as usually practised, influences the priorities and decision-making processes of researchers and research organisations.

The question of how science can be made more relevant to the needs of society is increasingly central in science-policy debate. Yet many of these discussions approach the question from the outside, as it were, preferring to leave the workings of science itself untouched; the focus is on the role of 'technology transfer', 'intermediaries' and 'dissemination'. In this article, I want to explore the inner workings of science by investigating its central decision-making tool: peer review.

Peer review plays a significant role in many of the key moments in science, as it is the main form of decision-making around: who receives money to do what science; who gets to publish in the scientific literature; and which individual scientists are selected and promoted within research institutions. Peer review is also the core tool used in various methods aimed at evaluating scientific institutions themselves: 'Peer review pervades science from beginning to end' (Science Media Centre undated but 2003).

Forces for relevance in science

A number of forces have in recent years been encouraging science in a 'relevant' direction, many of which will be familiar to readers of this journal. I do not wish to take up time and space rehearsing these in detail here, but briefly they have included:

- Pressures from science policy that have emphasised both the application of science and the need for accountability for public expenditure on science;
- Democratisation of science: critiques of the social impacts of science and technology have led to voices both within and outside the scientific community calling for more open discussion of scientific priorities and greater citizen participation in such discussions;
- New social and scientific challenges: the emergence of complex and controversial social problems characterised by high levels of uncertainty have created the challenge of 'post-normal' science.

Challenge of post-normal science

To take the last of these first, the appearance of a range of highly complex social challenges such as global environmental change have been accompanied by calls since at least the 1970s for science to lend more of its capabilities to addressing these challenges (for example, the 'Brooks' report: OECD 1971). Organisations such as the G8 and the OECD launch initiatives (OECD 2004; G8 group of nations 2003), and the worry that science is not doing all it could has also come from within the scientific community; Lubchenco's speech while she was President of the American Association for the Advancement of Science stands out in this regard (subsequently published in Science as: Lubchenco 1998).

However, this is not merely a matter of scientists becoming *more* involved in addressing social problems, important though this might be. It is also about the *character* of the scientific challenges changing, as suggested by the idea of post-normal science. Here, Ravetz and Funtowicz (see, among others, Funtowicz and Ravetz 1992; Funtowicz and Ravetz 2001) have suggested that science cannot simply apply its usual reductionist methods, but needs to grapple with the irreducibly complex, uncertain and socially controversial character of problems, echoing Weinberg's earlier term 'trans-science', where 'trans' indicates that the problems are beyond the capability of science to provide definitive answers (Weinberg 1972) and Commoner's still earlier thoughts about problems that are 'beyond the realm of science' (Commoner 1971 (1963): 120).

One of the main implications that Ravetz and Funtowicz draw from their work on post-normal science is the need for scientists to draw on 'extended peer communities', since the socially distributed character of the problems that scientists will be tackling means that many non-scientists will have substantive expertise on the topic under focus. Thus interactions with this extended peer community can help to ensure that researchers are: addressing the right questions; incorporating the knowledge of these non-academic experts in the analysis; and adequately testing the validity and practicality of any prescriptions researchers are proposing.

According to Ravetz and Funtowicz, traditional mechanisms of quality assurance are inadequate to cope with the needs of post-normal questions, although in saying this they do not clearly elaborate what would constitute adequate quality control in these circumstances. Instead, they seem to focus their recommendations on increasing consultation with 'extended peer communities' *during the research process*, leaving aside any consideration of how such consultation processes might affect the *design and selection* of research projects i.e. the peer-review process itself[1]. This is the question I wish to take forward in this article.

The defence of autonomy

The matter of decision-making in science has begun to emerge as a topic for debate, although few have managed to sustain their focus on the topic. Nevertheless, one of the implications of

[1] I should acknowledge that it is possible that I have missed something among these authors' substantial writings.

the pressures of accountability, democratisation and problem focus on science is that processes for setting the priorities of scientific work should become more open[2]. This is a matter of 'processual fairness', with the aim of arriving at proportionality of effort in relation to the urgency of the social problem (Rayner 1999; Lubchenco 1998). Where the case for research is made in terms of social problem solving, relevant stakeholders should be involved in setting the objectives (Shackley and Wynne 1996).

However, these perspectives are deeply contested in the scientific community, where the aspiration to autonomy means that it has long been widely believed that 'The freedom to choose his own problem is the scientist's most precious possession' (Commoner 1971 (1963): 59). This is no less than a battle for the integrity of science (ibid)[3]. For Ziman, this implies 'continued tension between internal scientific developments and external social demands' (Ziman 1994: 91). Quality criteria are both internal and external (ibid: 100-101), and some funding bodies have moved from peer review to the broader concept of 'merit' review 'indicating that they are taking account of non-specialist opinion on the relevance of the research to socio-economic problems' (ibid: 101).

For some this is worrying; the erosion of the internal scientific system of reputational control has been termed 'epistemic drift' (Elzinga 1985). Such authors are concerned to maintain a clear distinction between the social and cognitive aspects of science: 'It is important that the rejection of elitism, conservatism and academicism does not also lead to the rejection of internalist

---

[2] I use the term 'open' here in the sense used by the UK Royal Commission on Environmental Pollution in its report on 'Setting Environmental Standards' (RCEP 1998). This report called for greater *transparency* and *openness* in decision-making processes. For the RCEP, *transparency* means that people should be given access to the way in which decisions are made, including the information that underpins them, while *openness* means that decision-making processes should not simply be based on expert opinion, but should take account of a wider array of perspectives, especially of those people who stand to be affected by the decisions.

[3] Commoner is citing the Committee on Science in the Promotion of Human Welfare of the American Association for the Advancement of Science, which produced a report bearing this title in 1965. The report was published in *American Scientist* in June 1965; earlier discussions had been published in the pages of *Science* on 8th July 1960 and 29th December 1961. See also von Hippel and Primack's call for 'public interest science' in *Science* in 1972, 29th September, Vol 177, pages 1166-1171.

criteria and quality control mechanisms (e.g. peer review as such)' (Elzinga 1985: 197). Elzinga feared that any rise in the status of relevance criteria would lead to a 'skewing effect' or 'prostitution'[4] within the research community if 'strategic grants reach a level of 25% of the total resources allocated through the basic research councils' (Elzinga 1985: 202). Elzinga provided no explanation for choosing this threshold or the reasons why prostitution would come about at this level, but concludes that 'In science what is necessary is to find good people and let them do what they want' (ibid: 203). Here we see the tension between 'relevance' and 'autonomy' touched on by some eminent science-policy commentators (Rosenberg 1991).

Accountability

Science policy has also started to place greater emphasis on accountability, as part of a more general emphasis on accountability in what has been termed the 'audit society' (Power 1997). The growing costs of research and the need for public accountability bring 'demands for evaluation and for performance indicators to assure the government and the public that public money is being well spent' (Martin, Salter et al. 1996: 3; see also Ziman 1994).

The post-war belief that benefits will automatically emerge from 'arms-length' basic research is now being replaced by a more interventionist and utilitarian 'social contract' for science (Gibbons, Limoges et al. 1994; CVCP 1999). Under this contract, basic research is supported by governments 'but only if it generates rather more direct and specific benefits in the form of wealth creation and improvements to the quality of life' (Martin, Salter et al. 1996: 3).

As shown by Ruivo, many eminent authors have perceived the emergence of a new 'phase' in science policy since the early 1990s, in which publicly funded research has increasingly been expected to address social problems (Ruivo 1994; see also later contributions including OECD 1998; David and Dasgupta 1994).

[4] Elzinga uses these terms: see p.202.

Perhaps the best known of these contributions has been the Mode 1/Mode 2 hypothesis, which has proposed that research is increasingly taking place 'in the context of application'. Here, research is thought to be moving from 'Mode 1' – a 'traditional' style of knowledge production that is primarily based within universities and academic disciplines, and where 'problems' are primarily defined by researchers rather than others, to 'Mode 2'. The latter is a more trans-disciplinary, institutionally diverse, socially accountable and reflexive/interactive form of knowledge production that is focused on solving social problems defined outside the academy, or at least jointly between researchers and others (Gibbons, Limoges et al. 1994).

Gibbons and colleagues consider that one of the main differences between Mode 1 and Mode 2 is the way in which quality is assessed, both in deciding which research to fund and in evaluating research already conducted. Due to the more interactive nature of Mode 2 knowledge production, quality judgements need to be made on the basis of a wider set of criteria and involving a wider set of participants in making such judgements: scientific excellence is 'a necessary but not sufficient selection criterion for establishing research priorities' (Gibbons, Limoges et al. 1994: 65). Of particular importance is the relevance of the research to solving practical problems, as judged by practitioners, reminding us of the idea of 'extended peer communities' mentioned earlier.

Again, however, and similar to Funtowicz and Ravetz, these authors do not get beyond such general statements to suggest specific changes to the practice of scientific decision-making. In this paper, I hope to make a contribution in that direction. Before suggesting such changes, it is useful to review what we know of the history of discussions around peer review.

Peer review: a short history

Elsewhere, I have argued that various historical developments in the culture of science have served to create a broad agreement that the scientific community requires autonomy to be able to maximise its efficiency and creativity (Scott 2005). ‘Pure’ science is seen as the highest form of enquiry, an idea that started with the Greek founding fathers of science, when intellectual pursuits were made possible by the use of slave labour. This created the beginnings of the attitude that the pursuit of knowledge should be unsullied by considerations of practical utility (Tolstoy 1990; Stokes 1997; Flyvbjerg 2001).

The modern equivalent of pure science is ‘basic’ science, an idea that has been shown to be highly problematic when subjected to sustained critical scrutiny (Calvert 2001). This concept has had significant impacts on many of the assumptions that underlie traditional peer review, so is worth considering here briefly.

Definitions of ‘basic’ research

Calvert finds that basic research has been defined on the basis of three distinct sets of arguments, each of which is problematic: the intention behind the research; the epistemological characteristics of the research; and the autonomy of the researchers.

Under the intentional definition, basic research is research that is not aimed at practical outcomes. Such a definition is problematic since it is ‘attempting to define basic research in terms of a negative, as research which aims towards nothing’ (Calvert 2001: 11). There are also problems over who decides what is useful, and over what timescales, since much ‘basic’ research is justified on the basis that it will *eventually* become useful, even though it may lack specific uses at the time it is carried out.

The second, epistemological definition faces similar ambiguities. In some cases, such definitions state that basic research is the pursuit of knowledge about underlying properties, structures, theories and laws – it provides broadly applicable knowledge. This conflicts with another, this time reductionist version, which is that basic research provides knowledge at the level of constituent parts. Another epistemological definition relates to the unpredictability of the outcomes of research; a corollary of this is the frequent argument that interfering with basic research disrupts the creativity that ultimately leads to useful social outcomes (Nelson 1959; Polanyi 1962). However, unpredictability will partly depend on the way in which research is pursued; it also conflicts with other epistemological definitions since some types of research typically seen as basic, such as astronomy, are largely 'mundane and predictable' (Calvert 2001: 18).

Under the third definition, basic research is characterised by the *autonomy* of researchers to steer research. Autonomy may be highly valued by scientists, but it also creates problems as it implies that researchers, while receiving funds from government, are not accountable to it or may not be undertaking research seen by others to be socially relevant (Calvert 2001). Autonomy is underpinned by a system of peer review that determines scientific quality, and this is widely assumed to be central to the workings of science (Ziman 1994: 39; Elzinga 1985). However, this system tends to privilege epistemological criteria over others and it means that 'scientists are each other's judges'; non-scientists rarely have a role in peer review (Yearley 1988: 81; see also Frankel and Cave 1997; Redclift and Shove 1998). But the main problem with the autonomy definition is that 'it presumes that scientists can be immune to all 'external' interests, which sociology of science over the last thirty years has demonstrated is not the case' (Calvert 2001: 13).

Calvert concluded that despite, or even because of, these confusions and ambiguities, the idea of 'basic' science is useful to scientists and science policy-makers as a 'boundary concept' – an idea that is widely subscribed to but flexible enough to serve different needs and interests at different times.

Basic science into policy

We can see this in action when the idea of basic science was used by Vannevar Bush in his post-war advice to the US President in his report 'Science, the Endless Frontier'. Bush's aim was to win significant public funding for science on the back of its war-time contributions, at the same time as securing autonomy for scientists to make decisions over the direction of science. The usefully ambiguous idea of basic science provided cover for Bush's arguments, even though, as an engineer, he knew that much of the research that goes on in universities and other publicly-funded research institutions was, and would continue to be, directly aimed at addressing a range of social problems[5].

The problem with the success of the idea of basic science is that it has deeply affected the practice of peer review across science. It has served to widen the perceived gap between the value of those parts of science presented as 'basic' and those thought of as 'applied', exacerbating a problem already identified by C.P.Snow in his well-known observations about the divides within the scientific community. Snow is usually remembered for being critical of the divide between the sciences and the humanities, but a re-visit to his book shows that he was in fact just as critical of the divide between basic and applied science. As another author notes, there is 'more often than not, a hint of disavowal' when 'pure' scientists refer to applied science (Tolstoy 1990: ix).

The rise of the idea of 'basic' science has served to create a culture in which it is often simply assumed that scientific criteria should take primacy in decisions about scientific funding, publications and staff appointments, even when the aims of the research, for example, may be

[5] I give a fuller history of these developments in Scott 2004.

driven by concerns around practical utility rather than the search for scientific novelty. So what do we know about the criteria used in scientific decision-making?

Criteria in scientific decision-making

An early contribution to debates about the use of criteria in scientific choice was Weinberg's paper, in which he distinguished between criteria that were internal and external to science, and stated that 'it is not tenable to base our judgments entirely on internal criteria' (Weinberg 1963: 163). Weinberg's internal criteria were 'the quality of the researchers', and whether a field was 'ready for exploitation'. These internal criteria were thought unproblematic since they are the 'traditional' criteria used widely with apparent success.

Weinberg's external criteria were technological merit, scientific merit and social merit; Weinberg thought these less well defined[6]. He argued that scientific merit often consists in the degree of the potential contributions from research in one field to neighbouring fields, providing a justification for including specialists from these other fields in peer review processes. Weinberg says that in doing so, he is assuming that specialists outside a field can make useful contributions to judgements about research in that field (ibid: 166). This is interesting because it begins to break open the rather 'closed' logic, once promoted by writers such as Elzinga and often assumed by research funding bodies, that peer review should rely on specialists from the same field; if scientists from other fields can make valid contributions, practitioners might also play a useful role, as they also have their own specialist knowledge, reminding us of the arguments in favour of 'extending' peer communities.

With respect to 'social merit', Weinberg makes little progress other than to acknowledge that judgements on social merit involve values, that there are difficulties in identifying such values,

[6] It is interesting that Weinberg includes scientific merit among 'external' criteria, and that he states that 'we have given little thought to defining scientific merit in the broadest sense' (ibid: 164).

and even more difficulties in deciding whether a particular piece of research will further these values.

Weinberg devoted just one paragraph to this topic, perhaps because he found the problems too difficult to tackle[7]. However, even his brief observations indicated that given the above difficulties in identifying 'social merit', it would seem wise to encourage as much clarity in such judgements as possible. Weinberg did not seem to consider the possibility that the category of social merit might be elucidated by breaking it down into a range of sub criteria. This is after all the logic he employed in his article; in trying to inject clarity about scientific choice, he needed to identify the different criteria that might be used. Having gone this far, he seems to have decided not to pursue his own logic any further in relation to social merit.

Despite these weaknesses, commentators have acknowledged that Weinberg's article was useful in encouraging decision-makers to be clear about why they are supporting a particular piece of research, rather than simply relying on the somewhat 'mystical' and self-referring process of peer review (Yearley 1988: 80). Once we have opened the 'black box' of scientific choice, various problems come into view (Scott 2005): 'The fact that he {Weinberg} was unable to offer any way of weighting the various criteria against each other obviously limits the immediate practical utility of his work' (Yearley 1988: 80). Indeed, Weinberg acknowledged that judgements on the different types of merit are incommensurable[8].

However, there are other problems. Weinberg himself acknowledged the tensions between specialist and 'external' inputs to judgements about the potential worth of a piece of research.

---

[7] This would be ironic given that later in the paper he criticised social science on the basis that 'it is not clear to me that behavioural scientists, on the whole, see clearly how to attack the important problems of their sciences' (op cit: 169).

[8] That is, they cannot be measured on a common basis, posing limitations on comparisons.

He also stated that ‘The panel system is weak insofar as judge, jury, plaintiff and defendant are usually one and the same’ (ibid: 161)[9].

Persistent peer review

Since Weinberg’s early contribution, various commentators and committees of enquiry have grappled with a range of challenges in scientific decision-making. Most have encountered great difficulties and have failed to make significant progress. For me, one of the characteristics of peer review that indicates its power and structural role within science is its persistence.

Some commentators have started to analyse issues such as conflicts of interest in peer review (e.g. the influence of funding sources); an analysis in *Nature*, for example, revealed the wide variety of opinions and practices adopted by journal editors in this respect (Kolfschooten 2002).

In the UK, a national committee of enquiry into peer review (resulting in the ‘Boden’ report) noted various types of problems, concluding that these amount to ‘a salutary lesson for any with an uncritical enthusiasm for peer review’ (ABRC Working Group on Peer Review 1990: 10)[10]. It found a lack of clarity around peer-review processes, recommending greater transparency in, and frequent review of, the peer-review practices used by Research Councils.

In connection with our interests in this article – ‘relevant’ science – the most immediate problems are that peer review can present barriers to inter-disciplinary (see box) or risky research, although the Boden group ‘did not find substantial evidence’ that particular types of researcher, research approach or research topic were disadvantaged (op cit: 7; see also POST 2002). Nevertheless:

---

[9] Although by this we should be clear that he means (or should) that all of the players in peer review typically come from a specialist area within the academic community; clearly, scientists should not make peer review judgements on their own work.

[10] ABRC stands for the Advisory Board for the Research Councils, at that time one of the UK Government’s strategic sources of science advice.

> The following fates that afflict inter-disciplinary research and research in new fields were highlighted {in the consultation process}: 'shunted' between committees; sent to inappropriate referees; not given full attention because they are at the periphery of committees' interests; and rarely given highest grades because committees lack confidence in their judgments in the area (ibid: 32) [11].

To address this, the report recommended that: some funds should be ring-fenced for unorthodox proposals; that review processes should be flexible, for example to allow cross-membership on selection committees; and that applicants should 'flag' inter-disciplinary projects for special attention.

Box: The functions of disciplines

Scientific disciplines are a relatively new development in science: most have developed simultaneously with the institutionalisation of science (Kline 1995: 211). Most university science is organised on a disciplinary basis (ibid: 229); this is also linked to discipline-based teaching (Martin and Etzkowtiz 2000).

The exponential increase in scientific knowledge has made it necessary to have specific fields in which researchers can develop a workable knowledge of relevant information and methods. Disciplines are therefore a mechanism for simplifying research by dividing the world into manageable topics, methods and approaches.

But disciplines also embody research 'cultures' made up of assumptions, methods, and histories, to which new researchers are initiated. As well as demarcating *which* problems are to be tackled, disciplines therefore also help to define *how* problems are seen and tackled. In addition, the means of training (e.g. degrees) and academic communication (e.g. journals) are often controlled through disciplines, and processes of peer review, mostly within disciplinary boundaries, help to establish a hierarchy of the perceived quality of degrees and journals (Langenhove 1999). For these reasons, academic researchers often feel a stronger allegiance to the 'invisible colleges' represented by disciplines than to their own particular university (Kline 1995).

---

[11] According to the reports cited, other problems with peer review include fraud, bias (i.e. institutional bias, bias in favour of positive results, bias against women or young researchers), 'cloning' (i.e. favouring topics and researchers with characteristics similar to those of incumbent paradigms and researchers), and inefficiency.

Research funding agencies themselves often contend that as a result of its problem focus, 'relevant' science will need to be inter-disciplinary, and will often be risky in the sense that by addressing the full complexity of a problem, researchers run the risk of reaching inconclusive results (compare with the benefits of investigating a narrow set of phenomena via reductionist techniques).
End box

On the broader question of the social relevance of research, the Boden report acknowledged the importance of criteria such as timeliness and exploitability, in a broader process it termed 'merit review', and accepted that 'In theoretical literature, peer and merit review are separated: in practice, it is more difficult to distinguish the two' (ibid: 9). Unfortunately, the report largely failed to develop the implications from this finding.

Instead, the report resorted to the conclusion that 'excellence within one's specialism must be the pre-eminent criterion for committee membership' (ibid: 20, 47), although this needs to be balanced through a breadth criterion[12]. Most significantly, in a particularly circular and expedient piece of reasoning, the Boden inquiry concluded that while 'peer review does have problems both in principle and in practice' (ibid: 5) there was no practicable alternative to it because of the 'nearly universal agreement that peer review was the only appropriate and effective method of making decisions on the quality of basic research, and that it had the overwhelming support of the academic community' (ibid: 5).

Probably for this reason, none of the report's recommendations built on the acknowledgement that peer review and merit review cannot easily be separated, despite the report stating that 'Throughout we considered how peer and merit review…interplay closely' (ibid: 5). What is more, the conclusion relies on an invocation of 'basic research' as the situation in which peer

[12] 'Referees must be drawn as widely as possible' (ibid: 21)

review is indispensable; what of cases where basic research is not the aim and what, if we accept Calvert's analysis, if the idea of basic research is itself problematic?

The Boden Committee's evident difficulty in dealing with the broader demands made on peer review is not an isolated case. A later inquiry by the UK's Royal Society into peer review similarly saw no alternative to traditional peer review, despite making the same acknowledgement of the difficulties of separating peer and merit review (Royal Society 1995). Further, as we shall see shortly, various practical expressions of peer review such as the UK's Research Assessment Exercise also display confusions in this respect. Questions of social relevance are simply left hanging, with no conceptual apparatus for their inclusion into decision-making about priorities and resource allocation in science. Rather, the assumption seems to be that criteria such as timeliness, balance, relevance, and exploitability will be taken into account at a higher level:

> We have limited our consideration to decisions in which peer review plays the major role. In strategic and management decisions, or in assessing merit by criteria broader than the strictly scholarly, peer review can only play a part, and can properly be overridden by other considerations. Appeal to 'peer review' should not be used to legitimate decisions arrived at by other means (ibid: 11).

The question of how to include 'relevance' criteria in peer-review processes therefore seems to cause considerable difficulties even for national committees of enquiry chaired by eminent philosophers[13]. Evidence suggests that this problem is international in character, as supported by investigations into the organisation of relevant research in other countries (see for example Redclift and Shove 1998).

[13] Margaret Boden was Professor of Philosophy and Psychology at Sussex University, and a Fellow of the British Academy and member of the ABRC.

Quality-accountability mechanisms

Similar conclusions can be reached through an analysis of the procedures used to assess scientific performance, such as the Research Assessment Exercise (RAE), which is used by the UK Higher Education Funding Councils (HEFCs) to assess university departments. Such quality-accountability mechanisms can have strong effects on universities and on individual academics. For this reason, one observer has noted that if relevant research (in this case termed 'interactive') is to become mainstream, 'much will depend on the extent to which institutions of audit and control adopt 'interactive' as a measure of performance' (Woolgar 2000: 172).

Various concerns have been expressed about the design of such mechanisms[14], many revolving around their ability to reflect the complexity of academic activity through the use of a narrow set of indicators: 'The traditional criteria of scientific excellence are not adequate as stand-alone measures' (Royal Academy of Engineering 2000: 10). Evaluation of research 'tends to focus on the recent quantifiable outputs of the institution' (Geuna 1999: 29), rather than more intangible, informal and uncodifiable benefits that may characterise the outcomes of research processes, such as collaboration: 'Research selectivity exercises need to explicitly recognise the value of university-industry generated work, rather than simply rewarding effort via academic publications' (Charles, Pike et al. 1995).

These authors have suggested that an emphasis on short-term outcomes, and the publishing bias in favour of positive results, is likely to discourage risky or complex research. Inter-disciplinarity is disadvantaged by difficulties with publishing inter-disciplinary results, the widespread perception that inter-disciplinary journals are of a lower status, and by the fact that the RAE mostly works in disciplinary units.

---

[14] One senior administrator referred to the 'neo-Stalinist performance indicators' of the then University Grants Committee (the forerunner to HEFCE) (Hague 1991: 6).

For these reasons, such assessment mechanisms are seen by some to discourage relevant research. The UK House of Commons Science and Technology Select Committee's inquiry into the RAE was highly critical of its impacts on inter-disciplinary and collaborative research[15] (House of Commons Science and Technology Select Committee 2002). Academics themselves have commented on the discouraging effects the RAE has on 'third stream' activities[16] (Winfield 1996; Steele 1997).

Systemic failure

Several perspectives have been developed to explain the problems generated by quality-accountability mechanisms such as the RAE. As one analyst has found, audits work not on primary activities but on other systems of control (in the case of science, via the peer-review-validated literature as its proxy for quality and 'control'), leading to gaps between what is measured and what actually occurs in the organisation (Power 1994).

From a 'systems' perspective, such mechanisms are based on a model of policy-making that relies on reducing complex problems into separate components, and seeking to measure and control these components. But this approach leads to 'unintended consequences, alienation of professionals involved in delivery, and long-term failure to improve overall system performance' (Chapman 2002: 11).

As the Roberts review of the RAE itself acknowledged: 'all evaluation mechanisms distort the processes they purport to evaluate' (Roberts 2003: 3)[17]. The Roberts report cited a number of concerns about the RAE as the reason for the establishment of the review, including:

---

[15] Although the S&T Select Committee found it hard to identify witnesses who would go on record to criticise it (op.cit.).

[16] 'Third Stream' activities are those that go beyond the first two streams of work in universities: teaching and research. The term can be used to encompass a wide range of activities including community links, consultancy, policy advice and collaboration with the private sector. See Molas-Gallart, Salter et al 2002 for a fuller discussion.

[17] Power makes similar remarks in his work on audit systems.

- the need to fully recognise all aspects of excellence in research (such as pure intellectual quality, value added to professional practice, applicability, and impact within and beyond the research community);
- ability to recognise, or at least not discourage, enterprise activities;
- concern over the disciplinary basis of the RAE and its effects upon interdisciplinarity and multidisciplinarity (Roberts 2003: 4-5).

Despite this, not a single recommendation of the Roberts report mentions inter-disciplinarity or ‘enterprise activities’, and only one mentions dissemination beyond academia (ibid: 8). Therefore it seems that the criticisms of the RAE, despite being acknowledged as being important reasons for conducting this major review, received little attention in the review and in its recommendations. Again, ‘social relevance’ considerations are dropped with no explanation as to why. This observation supports the conclusion that peer review displays great persistence; indeed, we could say that it is ‘locked in’ in a similar way to many features of path-dependent technical systems, making ‘systems change’ hard to bring about (Hughes 1983).

This is likely to have consequences if we take seriously the remarks about audit mechanisms affecting the processes they are aiming to measure. Without being measured and rewarded, inter-disciplinary and interactive research will suffer active or passive discrimination. These measurement systems are therefore likely to reinforce existing incentives to focus on problems defined within disciplines, and to put effort into academic publishing rather than other activities. In this way, the incentives associated with the RAE are likely to work against the search for relevant science.

This tension was summarised in an exchange brought about by the following question in HEFCE's[18] 1997 consultation on the RAE: 'Should research assessment be concerned only with the question of research quality as opposed to, for example, its value for money or its relevance to wealth creation and the quality of life'. The response from the UK Economic and Social Research Council (ESRC) went as follows:

> The use of "as opposed to" implies that "research quality" is somehow separate from value-for-money or relevance. In many areas of the social sciences, value-for-money and relevance are essential aspects of research quality, not opposites. Quality as a concept must involve consideration of "fit for purpose"…taking this approach to research quality would undoubtedly be challenging…(ESRC 1997b: 1).

Here we see the multiple expectations that are placed on the research community, creating tensions between the aspirations of different funding agencies; this illustrates the multiple demands made of peer-review processes, as discussed next.

Multiple demands on peer review

A few of the scholars who have investigated peer review have begun to acknowledge the conflicting demands made on peer review. In a useful opening up of the multi-faceted character of peer review, Hackett concludes from a review of the literature that 'too often, discussions of peer review focus narrowly on technical matters…Analysts mesmerized by the mechanics of peer review…have often produced narrow empirical studies and skewed critiques' (Hackett 1997: 51-52).

Hackett builds on the work of Faust (Faust 1997), who contended that 'a "classical empiricist" philosophy of science is responsible for the belief that manuscripts and proposals can be evaluated according to standard, objective criteria (which in turn should give rise to reliable and unbiased judgments)' (Hackett 1997: 53)[19]. Hackett argues that, by contrast with this common

[18] Higher Education Funding Council for England
[19] This is Hackett's summary of Faust's position rather than a direct quotation.

assumption about the role of peer review – the need to reach a consensus and a 'right' decision – science is a pluralistic activity in which:

> Scientists are likely to disagree, and the problem for science and science policy is to make good use of this disagreement, not to treat it as aberrant or embarrassing… This is a very important point for those concerned with peer review because much energy has been expended on methods for 'homogenizing' peer ratings, whereas we may be better served by strategies for making better use of the inherent and valuable variability in those ratings (ibid: 54).

As a result, Hackett proposes that we need to reject over-simplified views of peer review for a more complex conceptualisation that acknowledges its 'diverse stakeholders, and the inconsistent demands they place on it' (ibid: 55). Hackett lists ten roles for peer review, from being a mechanism for improving research and publications, to being a 'counterweight to the drive for originality in science {as a} locus for enacting Kuhn's "essential tension" between tradition and innovation' (ibid: 55) to more abstract roles such as being 'an assertion of professional authority and autonomy that keeps the laity at bay' and 'a ceremony or ritual that affirms public trust in experts' (ibid: 56).

Serving all these roles involves a range of contradictions, and Hackett contends that peer review embodies tensions between five 'value pairs' – desirable properties that are in tension with each other:

- effectiveness (conducting thorough peer review) and efficiency (not taking too much time). Increases in effectiveness will require more work on the part of peer reviewers and will therefore impose greater costs, while high levels of efficiency will usually come at the expense of thoroughness.
- autonomy (making scientists accountable to scientific peers) and accountability (scrutiny by wider publics). Wider accountability might reduce autonomy; more autonomy implies less public accountability.

- responsiveness (helping identify new avenues of enquiry) and inertia (peer review is conservative because new ideas are judged against existing knowledge and approaches). This is the tension between tradition and originality.
- meritocracy (choosing 'the best' research) and fairness (e.g. for young researchers, or women). A poor paper by a respected academic may be published due to his or her reputation.
- reliability and validity – reliable criteria may be narrow and rigid, and thus may not produce the most valid results.

Although not completely critical of peer review, this analysis goes far in exposing the many possible tensions and opposed criteria that are expected to be embodied in peer review. Despite the difficulties involved in gaining research access to peer review processes, some studies have provided evidence to support these ideas (Chubin and Hackett 1990; Scott 2004). Others have pointed out that much ground-breaking work in e.g. economics, when submitted for publication, has initially been rejected in peer review (Gans 2000).

A way forward?

The discussion in the previous section has identified the danger that mainstream peer review is an overwhelmingly 'internalist' process aimed at identifying 'the best' scientific options. This is what I term 'autonomous mode' peer review, in which scientific opinions and criteria dominate to the exclusion of others (Scott 2005). The great tensions between tradition and originality within academic disciplines are played out through the peer-review process, and there is a tendency to downplay or completely ignore 'relevance' criteria. These are less well developed, and are rarely thought to have a central role in anything more specific than strategic decisions.

Analysis in earlier sections in any case suggested that, when the purpose is to fund research that is relevant to pressing social problems, there is a need to take into account a wider set of criteria than is normal in 'traditional' peer review. The search for relevance also means that peer-review processes should be more open, taking account of a wider set of views from the policy-makers and practitioners who may be substantively involved in addressing the challenge being researched. This final section makes some initial suggestions for taking forward a more eclectic and open form of peer review that might be more attuned to the needs of 'relevant science'.

The analysis given above points to a range of additional criteria related to the goal of relevance that need to be taken into account in peer review, including timeliness, urgency, significance, and exploitability. It also raises the profile of other criteria that may relate to the science but that may not normally receive much attention, such as balance within a research portfolio, and capacity creation. At the same time, any recommendation for changes to peer review needs to take into account the dissatisfaction with the burden it already imposes on research organisations and individual scientists[20].

Any suggested changes therefore need to be as clear, simple and practical as possible. Clarity will be assisted by making peer-review processes more transparent, so that judgements that are often based on implicit understandings are made more explicit. To help effect this, research organisations will need to provide explicit guidance to peer reviewers on the range of criteria on which they are seeking comment, especially if these are outside the normal range of criteria. This has already been attempted in some public research programmes (see for example Beemt and Pair 1991) and explored by others (Spaapen and Sylvain 1993; Meulen and Rip 1995; ESRC 1997b; Langenhove 1999) but also, particularly, in research foundations. The outcome orientation of the latter mean that they sometimes use a wider range of quality criteria and peer

[20] Such worries often arise in conversation with officials from organisations that fund research. Elzinga has argued that involving non-academics in steering research will lead to more onerous bureaucratic procedures, leading to 'inefficiencies' compared to usual peer review practices (Elzinga 1985). The ESRC's Open Door Scheme in the 1980s was apparently abandoned because of its administrative costs (Gill 1986; Caswill 1995).

reviewers (see for example the project selection methods used by the UK Alzheimer's Society at www.alzheimers.org.uk and the Joseph Rowntree Foundation at www.jrf.org.uk[21]).

Elsewhere, I have begun to sketch an approach that meets the demands of transparency, efficiency and flexibility, based on Stirling's development of the 'Multi-Criteria Mapping' method (Scott 2005; Stirling and Mayer 1999). Here, and in the spirit of debate, I end by outlining various exploratory 'relevance' and 'scientific' criteria for use in peer review. This list is exploratory because:

- it cannot be thought to be the stable conclusions of a sustained debate;
- different contributions have emphasised different aspects, but none have attempted a thorough synthesis to my knowledge;
- peer review is used in different contexts (project appraisal and evaluation, judging institutional performance, publishing etc), requiring different criteria.

Van den Beemt and Pair state that the following criteria have been used in a peer-review exercise in the Netherlands (Beemt and Pair 1991):

Scientific quality	**Utilisation potential**
Competence of the team Originality of the proposal Effectiveness of the research method Programme Time schedule Available infrastructure Costs Other aspects	Applicability in industry, society, technology or science Possibility of commercialising results Long-term contribution to technology advance Influence on the competitive power of Dutch industry Status of patents of Dutch industry Other aspects

[21] As at June 2005.

The Boden report acknowledged the need to produce ‘balance’ in peer review exercises, in terms of institution, age, gender, geography and subject (ABRC Working Group on Peer Review 1990: 19). It also stressed the desirability of avoiding ‘cloning’ i.e. research being formed in the image of research that has gone before. Thus there is a need for variety and breadth – essentially these are diversity criteria (ibid: 21). These echo some of Weinberg’s thoughts (Weinberg 1963: 161). The Boden report also indicated the ambiguity around peer review by stressing ‘the need to choose peers well’ (op.cit: 8), and the need to choose the ‘right’ peers (ibid: 51) (in the latter case the group seems to have been aware of the ambiguity of the concept as it put the word ‘right’ in inverted commas[22]).

Reliance on publications

Others have stated that quality-accountability mechanisms should not rely mainly on academic publication output as the measure of scientific productivity and value: ‘No single measure of quality can be used in isolation to present the true picture’ (Royal Academy of Engineering 2000: 6). The ESRC has stated that the RAE panels’ assessment of publication patterns should be ‘supplemented in the social sciences not just by other forms of peer review such as citations and journal hierarchies but by other appropriate steps to assess the quality of research in a wider sense, including value for money and relevance’ (ESRC 1997b: 5).

Other analysts, including those involved in the field of ‘bibliometrics’, have reached similar conclusions (see for example Clark 1985: section 3.3.6; Irvine and Martin 1983; Hicks and Crouch 1990). Crucially for our purposes, these authors acknowledge that:

> Citations and publications cannot be interpreted as straightforward indicators of scientific merit. There are four terms that are usually confused when describing the characteristics of scientific work that bibliometric indicators measure: quality, importance, impact, and citation (or publication) rates. ‘Quality’ describes how well the research has been done; it is a matter of judgment, and will therefore be evaluated differently by different people at different times. ‘Importance’ is the potential influence of a paper on a speciality: because

[22] It is interesting to note that on page 8 of the Boden report, the same page that it made the first ambiguous comment above, it dismissed the possibility of incorporating the concept of ‘fairness’ in peer review on the grounds of the ambiguity of its meaning.

> of the imperfections in scientific communication, the *actual* influence diverges from this potential and is called 'impact'. Each of these terms encompasses more contingent, external factors than the previous, including, for example, how well written the paper is, the eminence of its authors, their reading or referencing habits, and the size and dynamics of the field. Since further social factors intervene between impact and citation or publication rates, all one can measure are the latter, which are only partial measures of scientific impact' (Hicks and Crouch 1990: 28).

This analysis acknowledges the deeply social nature of an exercise that is commonly portrayed or taken to be technical in character. For our purposes, this is useful as it leaves the way open to ask whether the various aspects of user relevance could not similarly be analysed and taken account of in peer review. Indeed, similar to the Boden report (see p.12) the authors recommend that indicators should not be employed mechanistically but are rather:

> …most successful when individual scientists and institutions were involved in their construction and interpretation. Not only did a consultative approach produce an atmosphere of openness, reducing the hostility that inevitably follows any exclusively top-down approach, but it also enabled researchers to benefit from the review process themselves (Hicks and Crouch 1990: 28).

Diverse criteria, interaction, and diversity

Another point of departure in thinking about a more robust structure for peer review is the *diversity* of uses and communities of interest that publicly funded research serves (Royal Academy of Engineering 2000: 5). Thus relevance criteria might include the *range* and *intensity* of interaction and communication with different potential users (Wynne 1993: 9; Meulen and Rip 1995; Royal Academy of Engineering 2000). Relevance criteria could also include potential social usefulness, acceptability (of outcomes), legitimacy (of process), even expediency. Langenhove suggested relevance criteria around completeness, reliability, accessibility and timeliness (Langenhove 1999).

The Royal Academy of Engineering also suggested 'independence' as a criterion, which it defined as researchers' freedom from financial or political control (ibid: 13). It developed its own scheme for judging the value of research work, based on a broader set of criteria than academic publications alone. This was in reaction to the Research Assessment Exercise, which it saw as unduly narrow; even though the RAE ‘is recognised as amongst the most developed approaches available…it has shortcomings mainly in relation to the assessment of exploitation mechanisms’ (ibid: 6).

Other reports have identified further criteria, both scientific and otherwise (see for example Weinberg 1963; Wynne 1993; Royal Society 1995; POST 2002). Other *scientific* criteria relate to the potential scientific importance of a research proposal to its own or other fields of enquiry; the scale or diversity of the activity; the level of inter-disciplinarity; the level of ‘risk’ involved (e.g. potential for scientific outcomes); the value of investing in alternative lines of inquiry; potential to challenge established ideas; and potential for theoretical or methodological development.

In addition, *organisational* criteria might relate to: potential to create national scientific pre-eminence in a field; potential to support international collaboration; the need to invest in focused centres of excellence, or, conversely, to assure strength in diversity by investing in a variety of different scientific groups; the wish to support ‘new’ universities in establishing research excellence.

Conclusions

In conclusion, peer review potentially includes a range of criteria – scientific, organisational, and relating to social relevance – all of which are largely concealed in the current policy of treating peer review as a ‘black box’ – a tool for the exclusive use of the scientific community that is thought best left to work autonomously.

It would be premature to suggest a definitive list of criteria for use in peer review, but the above discussion summarises various contributions to suggest a range of criteria. The employment of a more structured approach to the criteria used in peer review might offer the chance to increase the levels of transparency and openness in peer review.

A more plural approach to the criteria used in peer review would be an outcome that might lead to significant advantages for those involved in conducting, and being the subjects of, peer review. It might also help to ensure that peer review is not seen as an opaque affair through which scientists escape scrutiny and accountability.

References

ABRC Working Group on Peer Review (1990). Peer Review (the 'Boden' report): A report to the ABRC. London, Advisory Board for the Research Councils: 82.

Beemt, F. v. d. and C. l. Pair (1991). "Grading the grain: consistent evaluation of research proposals", in Research Evaluation **1**(1): 3-10.

Calvert, J. (2001). Goodbye Blue Skies? The concept of basic research and its role in a changing funding environment. SPRU. Brighton, University of Sussex**:** also at www.centres.ex.ac.uk/egenis/staff/calvert/publications.php.

Caswill, C. (1995). Interactive social science workshop. Swindon, ESRC.

Chapman, J. (2002). System Failure: Why governments must learn to think differently. London, Demos.

Charles, D., A. Pike and J. Howells (1995). The impact of international collaboration on UK university-industry links. European context of UK science policy programme: programme briefing**:** No. 6.

Chubin, D. E. and E. J. Hackett (1990). Peerless science: peer review and US science policy. New York, State University of New York Press.

Clark, N. (1985). The political economy of science and technology. Oxford, Basil Blackwell.

Commoner, B. (1971 (1963)). Science and survival. London, Ballantine books.

CVCP (1999). Higher education in the 21st Century: some possible futures. London, Committee of Vice Chancellors and Principals: 15.

David, P. A. and P. Dasgupta (1994). "Toward a new economics of science", in Research Policy **23**(5): 487-521.

Elzinga, A. (1985). "Research, Bureaucracy and the Drift of Epistemic Criteria" in The University Research System: The Public Policies of the Home of Scientists edited by B. Wittrock and A. Elzinga. Stockholm, Almqvist & Wiksell International.

ESRC (1997b). ESRC response to Research Assessment consultation (RAE 2/97). Swindon, ESRC: 11.

Faust, D. (1997). "Research on the Process of Journal Review Re-Viewed", in Social Epistemology.

Flyvbjerg, B. (2001). Making social science matter: Why social inquiry fails and how it can succeed again. Cambridge, Cambridge University Press.

Frankel, M. S. and J. Cave, Eds. (1997). Evaluating science and scientists. Budapest/Oxford, Central European University Press/OUP.

Funtowicz, S. and J. Ravetz (2001). "Post-normal science. Science and Governance under conditions of complexity" in Interdisciplinarity in Technology Assessment edited by M. Decker, Springer**:** 15-24.

Funtowicz, S. O. and J. R. Ravetz (1992). "The good, the true and the post-modern", in Futures **24**(10): 963-976.

G8 group of nations (2003). G8 - Evian Summit 2003 - Chair's Summary and Action Plan on Science and Technology for Sustainable Development. G8 groups of nations, see http://www.g8.fr/evian/english/home.html.

Gans, J., Ed. (2000). Publishing Economics: An analysis of the academic journal market in economics. Cheltenham UK and Massachusetts, USA, Edward Elgar.

Geuna, A. (1999). The economics of knowledge production. Funding and the structure of university research. Cheltenham, Edward Elgar.
Gibbons, M., C. Limoges, H. Nowotny, S. Schwartzman, P. Scott and M. Trow (1994). The New Production of Knowledge. London, Sage.
Gill, J. (1986). "Research as action: an experiment in utilizing the social sciences" in The Use and Abuse of Social Science edited by F. Heller. London and New Delhi, Sage**:** 99-122.
Hackett, E. J. (1997). "Peer review in science and science policy" in Evaluating science and scientists edited by M. S. Frankel and J. Cave. Budapest/Oxford, Central European University Press/OUP.
Hague, D. (1991). "Can scientists manage science?" in The management of science edited by D. Hague. Basingstoke and London, Macmillan.
Hicks, D. and D. Crouch (1990). "Can bibliometrics measure up?", in Physics World **3**(9): 27-28.
House of Commons Science and Technology Select Committee (2002). The Research Assessment Exercise. London, House of Commons, HMSO: 175.
Hughes, T. P. (1983). Networks of Power: Electrification in Western Society 1880-1930. Baltimore, John Hopkins University Press.
Irvine, J. and B. Martin (1983). "Assessing basic research: some partial indicators of scientific progress in radio astronomy", in Research Policy **12**(2): 61-90.
Kline, S. (1995). Conceptual foundations for multidisciplinary thinking. Stanford, CA, Stanford University Press.
Kolfschooten, F. v. (2002). "Can you believe what you read?", in Nature **416**: 360-363.
Langenhove, L. V. (1999). "Re-thinking the social sciences? A point of view" in The social sciences at a turning point? edited. Paris, Organisation for Economic Cooperation and Development**:** 43-51.
Lubchenco, J. (1998). "Entering the century of the environment: a new social contract for science, Presidential address to the American Association for the Advancement of Science, 15th February 1997", in Science **279**(23 January 1998): 491-497.
Martin, B. and H. Etzkowtiz (2000). "The origin and evolution of the university species", in VEST: Swedish journal for science and technology studies **13**(3-4): 9-34.
Martin, B., A. Salter, K. Pavitt, J. Senker, M. Sharp and N. v. Tunzelmann (1996). The Relationship Between Publicly Funded Basic Research and Economic Performance. A SPRU Review for the UK Treasury. Brighton, SPRU - Science and Technology Policy Research.
Meulen, B. v. d. and A. Rip (1995). Assessing societal quality of research in environmental sciences. A report to the Consultative committee of Advisory Research Councils and the Advisory Council for Research on Environment and Nature. Twente, University of Twente Centre for Studies of Science, Technology and Society: 34.
Nelson, R. (1959). "The simple economics of basic research", in Journal of Political Economy **6**: 297-306.
OECD (1971). Science, growth and society: a new perspective. The 'Brooks report'. Paris, OECD: 113.
OECD (1998). University research in transition. Paris, Organisation for Economic Cooperation and Development: 103.

OECD (2004). Science, Technology and Innovation for the 21st Century. Meeting of the OECD Committee for Scientific and Technological Policy at Ministerial Level, 29-30 January 2004 - Final Communiqué. Paris, Organisation for Economic Cooperation and Development.
Polanyi, M. (1962). "The republic of science: Its political and economic theory", in Minerva **1**: 54-73.
POST (2002). Peer Review. London, Parliamentary Office of Science and Technology: 4. POSTnote 182.
Power, M. (1994). The Audit Explosion. London, Demos.
Power, M. (1997). The Audit Society: rituals of verification. Oxford, Oxford University Press.
Rayner, S. (1999). The case for interactive social science - opening remarks. Interactive Social Science -- the production and consumption of social science research, Brighton, Association of Learning Societies in the Social Sciences.
RCEP (1998). Setting Environmental Standards. London, Royal Commission on Environmental Pollution,. 21st Report.
Redclift, M. and E. Shove (1998). Programming Social Environmental Research and Research Social Environmental Programme. A report to the ESRC. Swindon, ESRC: 25.
Roberts, S. G. (2003). Review of Research Assessment: report issued for consultation May 2003. Executive Summary. RA Review: 20.
Rosenberg, N. (1991). "Critical issues in science policy research", in Science and Public Policy **18**(6): 335-346.
Royal Academy of Engineering (2000). Measuring Excellence in Engineering Research. A statement identifying the characteristics of engineering research and a proposed methodology for assessing excellence. London: 28.
Royal Society (1995). Peer review: an assessment of recent developments. London, Royal Society: 8.
Ruivo, B. (1994). "'Phases' or 'paradigms' of science policy?", in Science and Public Policy **21**(3): 157-164.
Science Media Centre (undated but 2003). Peer review in a nutshell: a guide for your news interview. London, Science media centre: 6.
Scott, A. (2004). 'Relevant' social science? The Case of Global Environmental Change Research in UK Universities. SPRU (Science and Technology Policy Research). Brighton, University of Sussex**:** 270.
Scott, A. (2005). "Peer review: opening the black box", in Submitted to The Lancet; SPRU Electronic Working Paper series at www.sussex.ac.uk/spru(***).
Shackley, S. and B. Wynne (1996). "Representing Uncertainty in Global Climate Change Science and Policy: Boundary-Ordering Devices and Authority", in Science, Technology and Human Values **21**(3): 275-302.
Spaapen, J. and C. Sylvain (1993). "Assessing the value of research for society", in Research Evaluation **3**(2): 117-126.
Steele, F. (1997). Recognise that research transcends the RAE grade. Research Fortnight. **29 January**.
Stirling, A. and S. Mayer (1999). Rethinking Risk. A pilot multi-criteria mapping of a genetically modified crop in agricultural systems in the UK. Brighton, SPRU - Science and Technology Policy Research: 80.

Stokes, D. E. (1997). Pasteur's Quadrant: Basic science and technological innovation. Washington DC, Brookings Institution Press.
Tolstoy, I. (1990). The Knowledge and the Power: Reflections on the history of science. Edinburgh, Canongate.
Weinberg, A. (1963). "Criteria for Scientific Choice", in Minerva **1**(2): 159-171.
Weinberg, A. (1972). "Science and trans-science", in Minerva **10**(2): 209-222.
Winfield, A. (1996). The academic's dilemma: industry versus the RAE. Research fortnight.
Woolgar, S. (2000). "Social basis of interactive social science", in Science and Public Policy **27**(3): 165-173.
Wynne, B. (1993). Dissemination and the ESRC GECP: some issues for discussion at the GECP/CSEC workshop on dissemination. Lancaster, CSEC, Lancaster University: 10.
Yearley, S. (1988). Science, technology and social change. London, Unwin Hyman.
Ziman, J. (1994). Prometheus Bound: Science in a dynamic steady state. Cambridge, Cambridge University Press.