

Peer Review: the challenges for the humanities and social sciences

A British Academy Report

September 2007



The British Academy

THE NATIONAL ACADEMY FOR THE HUMANITIES AND SOCIAL SCIENCES

*The British Academy
10 Carlton House Terrace
London
SW1Y 5AH*

*Telephone: 020 7969 5200
Fax: 020 7969 5300
Email: chiefexec@britac.ac.uk*

Web site: www.britac.ac.uk

© The British Academy 2007

*Designed and produced
by Quattro
+44 (0)20 7766 5225*

Peer Review: the challenges for the humanities and social sciences

A British Academy Report

September 2007



British Academy Reports and Policy Statements

The British Academy is the UK academy for the humanities and social sciences. The Academy publishes reports on issues of particular importance to these disciplines, and to the formation of public policy that bears on them.

Examples of recent Academy reports and statements are shown below. These are all available online from www.britac.ac.uk/reports/

Peer Review: the challenges for the humanities and social sciences, A British Academy Report, September 2007

Dearing Review of Language Policy - Response to the Interim Consultation Report, February 2007

Dearing Review of Language Policy - Response to the Announcement of the Review's Establishment, November 2006

Response to the AHRC Review of Postgraduate Funding, November 2006

Response to the DfES Consultation on the Reform of Higher Education Research Assessment and Funding, October 2006

Copyright and Research in the Humanities and Social Sciences, A British Academy Report, September 2006

Response to a Consultation on the Draft Research Agenda for Theme 8 'Socio-Economic Sciences and the Humanities' in the 7th Community RTD Framework Programme, 2007-2013, July 2006

Response to the British Library's Content Strategy, June 2006

Response to the Gowers Review of Intellectual Property, April 2006

Response to the AHRC's Consultation - A Strategy for Supporting and Sustaining High Quality Research in the UK's Museums, Galleries, Libraries and Archives, March 2006

Response to HEFCE's Review of the Teaching Funding Method, June 2006

Comments on HEFCE's Draft Strategic Plan for 2006-11, January 2006

Response to the AHRC's Consultation on Doctoral Research in the Arts and Humanities, November 2005

Foreword

Peer review by expert assessors is used to judge the quality of research and writing in many disciplines, and is of great importance to the humanities and social sciences. It provides expert judgements that can form a basis for decisions whether to publish works of scholarship, or to fund proposed research. Such judgements also provide feedback to researchers and writers as they move towards completion of a work. The high international reputation of UK research in the humanities and social sciences is based on the rigour with which peer review is used, and the care with which practical decisions to fund or publish are based on it.

The British Academy has undertaken this review in the hope that it will provide policy makers and others with useful analysis of current practices of peer review in the humanities and social sciences, of some of the problems and difficulties these practices face, and of ways in which they might be improved. Some recent discussions of peer review have contrasted it with 'metrics' (sometimes said to be more reliable, more objective or cheaper to administer); others have raised questions about the increasing burden and cost of peer review; yet others have commented on the effects of using peer review for second-order purposes, such as assessing university departments and research groups. This report provides both evidence and analysis that bears on all of these concerns.

The Academy is grateful to Professor Albert Weale, who chaired the Working Group that oversaw the Review, to other members of the Group and to the members of Academy staff who have contributed significantly to this report.

Baroness Onora O'Neill, CBE, PBA, F Med Sci, Hon FRS
President of the British Academy

Contents

British Academy Reports and Policy Statements	ii
Foreword	iii
Preface	vii
Executive Summary and Recommendations	
Introduction	ix
Recommendations	xi
Chapter 1: Peer Review - the Changing Context	
1.1 The Importance of Peer Review	1
1.2 Primary Peer Review: the Practice	2
1.3 The Changing Context	4
1.4 Past Studies	6
1.5 Principal Questions	6
Chapter 2: Peer Review for Publications	
2.1 Introduction	8
2.2 What is Journal Peer Review?	8
2.3 Is Peer Review Needed?	11
2.4 Principles of Good Practice	12
2.5 Conflicts of Interest	15
2.6 Scholarly Malpractice	15
2.7 Conclusions	17
Chapter 3: Peer Review and Grants	
3.1 Introduction	18
3.2 AHRC and ESRC Peer Review Processes	18
3.3 British Academy Peer Review Processes	19
3.4 How is Peer Review Regarded by the Research Community?	20
3.5 Examples of Good Practice in Quality Assessment	21
3.6 Transparency of Peer Review Practices	22
3.7 Conclusions	23
Chapter 4: Training and Prior Preparation for Peer Review Activities	
4.1 Introduction	24
4.2 Training Available to Postgraduates	24
4.3 Training in Grant Awards	25
4.4 The Contribution of Non-Academic Experts to Peer Review	25
4.5 Conclusions and Recommendations	26

Chapter 5: The Costs of Peer Review

5.1	Motivations and Incentives	27
5.2	Journal Review and Full Economic Cost	28
5.3	The Costs of Peer Review for Grants	29
5.4	Conclusions and Recommendations	30

Chapter 6: Metrics and Peer Review

6.1	Introduction	31
6.2	What is Meant by Metrics and Metrics-based Research Assessment	31
6.3	The Citations Approach	32
6.4	Monographs and Metrics	33
6.5	The European Reference Index for the Humanities (ERIH)	34
6.6	Conclusions and Recommendations	37

Chapter 7: Peer Review and Innovation

7.1	Introduction	38
7.2	Innovation and Publication	38
7.3	Innovation and Grants	39
7.4	Knowledge Transfer and Impacts	40
7.5	Conclusions and recommendations	41

Appendices

Appendix 1 - How the Review was Conducted	45
Appendix 2 - List of Background Documents and Sources of Information	47

Preface

“Judge not that ye be not judged” is an impossible principle to follow in academic life. The selection of work for publication, the awarding of grants and the funding of university departments for research all depend upon the exercise of judgement. The distinctive feature of this academic judgement, however, is that it is reciprocal. Its guiding motto is: judge only if you in turn are prepared to be judged. This is one way of expressing the principle of peer review.

The present report contains the deliberations of the Working Group established by the British Academy to investigate the operation of peer review in the humanities and social sciences. As we worked through the issues and met editors, those active in awarding research grants and others involved in peer review, it was striking how much support the principle of peer review secured. This is not to say that those who gave us evidence thought that there were no improvements to be made, but the improvements suggested were those to a common purpose, namely the maintenance and enhancement of the quality of academic work.

On behalf of the Working Group, I should like to express our gratitude to all those who supplied us with material and evidence. Even when their words are not directly used in this report, they can be assured that they helped always to shape our thoughts. Drafts of the report were read - peer reviewed - by a number of people listed in the Appendix. They saved us from a number of mistakes.

Finally, I should also like to express the gratitude of the Working Group to our secretary, Vivienne Hurley, whose remarkable ability to record our work as it progressed facilitated our thinking, and whose good humour kept us going through long meetings.

We hope this report will be of use in stimulating reflection and debate on the way that peer review does, and can in future, operate to the benefit of scholarship in the humanities and social sciences.

Professor Albert Weale, FBA
Chairman of the Academy Working Group

Executive Summary and Recommendations

Introduction

Peer review is the practice by which the worth of research is evaluated by those with demonstrated competence to make a judgement. It is the traditional means by which research quality is guaranteed in academic studies. The Working Group was established to examine how the practice of peer review functioned at present in a context in which its scope was expanding beyond its traditional primary focus on individual publications and grants to encompass broader evaluations of, say, the research performance of departments. These kinds of evaluations, of which the most important in the UK has been the Research Assessment Exercise (RAE), typically involve assessors examining in a summary way publications and grant-aided research that have already been peer reviewed, so they can be thought of as second-order or secondary peer review.

The Working Group met five times and held fact-finding and opinion-sounding meetings with a number of interested groups and individuals. Throughout our deliberations, we sought to pay particular attention to issues specific to the humanities and social sciences, whilst being aware that a number of issues are general to any form of academic enquiry.

Conclusions

Recent developments raise questions about the way in which peer review is conducted and will be conducted in the future. Some of these developments are positive (such as advances in information technology, which speed up the process and also make it easier to use international referees). Others pose particular challenges (such as the increase in the volume of submissions both for grants and for journals), which add to the burdens of peer review. The proposed development of metrics to play a more prominent role in the assessment of the research performance of departments (RAE) interacts in complex ways with the practice of peer review, which we discuss in the report.

Peer review has its critics, who allege that it is costly, time-consuming and biased against innovation. None of these criticisms is entirely without force, but the Working Group concluded that there were no better alternatives and that often the criticisms were directed at deficiencies of practice rather than the principle of peer review. Peer review is both a mechanism of selection - only those grants and publications are favoured that are positively judged by peers - and a force making for enhancement. Work is better as a result of peer review. Importantly, it retains widespread and deep support among members of the academic community.

Peer review in practice takes a wide variety of forms, reflecting the diversity of subject matter and approaches in humanities and social science research. There is no one model that all should follow. We agree with RCUK that peer review is not in fact a single process, but rather a flexible set of mechanisms¹.

This variety of practice is important in relation to publication. There are many different models of peer review used. It is a considerable merit of the way in which the peer review works in journal publications that there is not one single model of good practice that all should follow, but instead decentralised diversity. Nevertheless, there are principles that good peer review should follow. These include timeliness, transparency and verifiability. These principles cannot guarantee the identification of the best quality work on a fair basis, but without them quality and fairness will suffer.

In the case of grants peer review remains essential if good work is to be identified. In a situation in which applicants have few alternatives to funding, it is important that funding bodies uphold the

¹ RCUK Response to the Project Report and Consultation on the Efficiency and Value-for-Money of Peer Review, June 2007

integrity of their peer review processes. It is also important that they find ways of responding to the innovative and the risky.

Recommendations

Our conclusions lead to the recommendations below, which are shown under the following headings:

- training;
- the costs of peer review;
- peer review and metrics;
- peer review and innovation.

The practice of peer review is complex, so that many of our recommendations under specific individual headings also relate to those under other headings. We have not prioritised our recommendations as they are addressed to various bodies whose individual concerns are different.

Training

Because peer review relies upon professional norms, it is important that these are reinforced by appropriate training.

R1 Training guidelines should be amended so that explicit attention is given to the need to train postgraduates in the norms and conventions of peer review. Training should also be made available to early career track postdoctoral researchers, which could be an effective use of some of the 'Roberts' monies'.

For Research Councils UK (RCUK), Arts and Humanities Research Council (AHRC) and Economic and Social Research Council (ESRC), other HSS research funders, universities and their representative bodies

We were impressed by the AHRC induction days, which enable panellists and reviewers to train alongside each other. The AHRC panellists we interviewed said that reviewers had a much better understanding of what was expected of them, and that the quality of their reports had improved markedly. We encourage other research funders to look at AHRC practice to see if there are any lessons that they can draw from it.

R2 Recognise the importance of training to ensure the competence and propriety of referees. We commend the AHRC programme of induction and training for its panellists and members of its Peer Review College.

For humanities and social science (HSS) research funders

The Research Councils are under pressure to employ non-academic peer reviewers. In many circumstances this is to be welcomed, but the practice highlights the need for training and clear communication.

R3 Give further consideration to the way in which non-academic expert users contribute to research council peer review assessments, with particular emphasis on the training and support that they receive.

For AHRC, ESRC and other research funders

The Costs of Peer Review

Surveys and other evidence have shown that there are various reasons why academics participate

in peer review. Not all motivations are altruistic, and there is no reason why they should be. However, a central element, without which the peer review system would not exist, is the professional commitment to contribute to the academic public good. There is no practical way, particularly in the humanities and social sciences, in which the full economic cost (FEC) of peer review activity for the purpose of journal publication can be recovered. This reality needs to be recognised within the funding of research and universities.

By a similar token, the application of FEC would make the funding of 'open access' journals in the humanities and social sciences impossible.

The significant contribution made by peer review to the health, well-being and development of academic disciplines needs to be better understood and recognised. Peer review has undoubtedly helped to underpin the success of UK research in the humanities and social sciences. The importance of these activities should be better reflected in research assessment exercises. Those responsible for the management of universities and research institutes need to ensure that they have processes in place to encourage and reward peer review activity, recognising that the benefits to the institution are indirect rather than direct.

R4 Each university in receipt of public funds should accept an obligation to encourage its researchers to engage in these activities, recognising that peer review is an essential part of the fabric of academic life - the costs of which are met by the funds allocated by the Funding Councils to support research infrastructure.

For University Vice-Chancellors and Universities UK (UUK)

Peer review will always impose costs, not least upon grant applicants, and the financial burdens of peer review can be measured in different ways. Any method that simply looks at the time-costs of peer review relative to size of awards is bound to cause problems in the humanities and social sciences, where awards are typically smaller than in medicine and the natural sciences. Policy makers need to be aware of these differences, and we commend RCUK² for taking note of them. Research Councils need to ensure that their standards of peer review, whilst robust in respect of quality, are proportionate in terms of the burdens they impose.

R5 Develop a more sophisticated understanding of the costs of peer review, and maintain efforts to ensure that the peer review processes are proportionate to the scale of award.

For RCUK/Funding Councils/relevant Government Departments

Metrics and peer review

Metrics and peer review are related in complex ways. Metrics are typically summary measures of peer reviewed activities, and the use of metrics sets up incentives to change behaviour in ways that affect peer review.

R6 Care should be taken to ensure that any metrics employed reflect the distinctive nature of the humanities and social sciences research and do not have an adverse affect on the quality of the work that they are seeking to measure. HEFCE should commission an investigation to explore whether there is scope to modify the Thomson Scientific indices to accommodate the special features of humanities and social science research. Until this work is completed and assessed, metrics should remain an adjunct to the research assessment (RAE) panel peer review process rather than a substitute.

For Higher Education Funding Council for England (HEFCE)/RCUK/relevant Government Departments

Difficulties with existing metrics have led to attempts to develop alternative approaches, notably the European Reference Index for the Humanities (ERIH) that involves the ranking of journals. However, the Working Group identified a number of methodological difficulties with the ERIH approach, relating to reliability and validity, some of which were fundamental.

R7 The experience of the European Reference Index for the Humanities (ERIH) shows how difficult the task is, and we conclude that the ERIH does not at present represent a reliable way in which summary measures of peer reviewed publications can be constructed.

For Higher Education Funding Council for England (HEFCE)/RCUK/relevant Government Departments

Peer review and innovation

The variety of practices in the conduct of peer review among journal editors in the humanities and social sciences is a strength, not a weakness.

R8 Uphold the editorial independence of journal editors.

For publishers and learned associations

Research funders should take pains to avoid a mechanistic approach in their decision-making processes for the award of research grants in the humanities and social sciences. One example of good practice in this regard is the AHRC's structure of subject-specific panels. As with journal editors, AHRC panels are able to reach independent assessments of quality and can treat reviewers as advisers rather than judges. Each proposal is considered on a case by case basis, avoiding a formulaic averaging of grades, so one low mark will not automatically rule out a proposal. Similarly, the ESRC processes for its standard grants are designed with safeguards to handle proposals receiving divergent reviews and marks. These approaches also help to ensure that intellectually innovative proposals, where there is likely to be a marked contrast in the views expressed by peer reviewers, can be assessed on their merits.

R9 Consider grant proposals on a case by case basis, taking pains to ensure that award decisions are not made by the application of formulae.

For HSS research funders

Issues of knowledge transfer and impact play an important role in public policy, and are likely to become more, not less, important over time. Quality should not be sacrificed in favour of relevance and impact. Applied research ought to meet the same standards of research design, sample selection and evidential inference that applies to any sort of work (allowing for the practical difficulties of conducting applied research). Indeed, if research is being used by policy makers to take decisions on matters that have a direct effect on the quality of citizens' lives, the standards ought to be as high as possible.

Similarly, novelty cannot be regarded as a substitute for quality.

R10 Ensure that considerations of applicability and relevance do not compromise judgements of quality.

For HSS research funders

We commend the AHRC for seeking to ensure that innovative proposals are dealt with through funds that explicitly recognise a degree of risk. Other research funders for the humanities and social sciences may wish to consider adopting the same practice.

R11 Set aside funds for risky, speculative projects.

For HSS research funders

We commend a balance between speculative, high-risk research and safer work that develops ideas already well established in response mode funding.

R12 Ensure that there is a healthy balance between strategic funding and responsive mode projects.

For HSS research funders

The use of small pump priming grants for innovative research in the humanities and social sciences can be very effective, especially those that are targeted at early career track researchers and those with novel approaches. A number of universities offer small pump priming grants to their staff with these aims in mind, and often support these grants by drawing on their endowment funds. This seems to be a sensible way to use these funds, and there may be scope for the Government to consider ways in which it can encourage endowments of this kind within universities.

R13 Encourage endowments within universities to support small grants for innovative, high risk research.

For UK Government, UK universities, UUK and other representative bodies for universities

The selection of topics for strategic initiatives is not straightforward, and should be subject to robust peer review procedures.

R14 Ensure that the process of selecting topics for strategic initiatives is also subject to peer review.

For HSS research funders

Chapter 1: Peer Review - the Changing Context

“One suspects that peer review is a bit like democracy - a bad system but the best one possible. It seems to be one that takes different forms in different (scientific) cultures and can be tweaked to improve its operation.”³

1.1 The Importance of Peer Review

What happens when a researcher submits an article to a journal or a book manuscript to a publisher? Or what happens when a researcher applies to a funder for a grant to carry out a piece of work? The answer in most cases - and certainly in all important cases - is that the proposed publication or the grant application will be sent out for peer review. Peer review is the practice by which research is assessed by experts in the appropriate field of study for its worth and value.

Peer review in the field of journal publications has been traced back to the late seventeenth century, when in 1665 the *Journal des Sçavans* in France and the *Philosophical Transactions of the Royal Society* in England introduced the practice of asking the opinion of those qualified to judge whether a report or an article should be published.⁴ As Zuckerman and Merton note, research peer review is but one instance of a general phenomenon by which judges are charged with evaluating role-performance in a social system. Arts critics and sports coaches have a similar status. In any social system in which the assessment of performance is an important element of decision making, there is the role of reviewer to be filled.⁵

Within the world of research, the *primary* role of peer review has been to assess individual publications and grant applications. However, as the assessment of performance has become more extensive in higher education and research policy making, peer review has been used more widely, and now includes the following:

- Evaluations of the research performance of departments and similar units (research assessment exercise - RAE).
- Departmental reviews in universities.
- Evaluations of research centres by Research Councils.
- Assessments of journals, for example the European Reference Index for the Humanities (ERIH).
- Reviews of disciplines by the Research Councils.
- International bench-marking reviews of disciplines of the sort recently carried out by the Economic and Social Research Council (ESRC) in conjunction with relevant professional bodies.

Because this more extensive list of activities typically involves assessors examining in a summary way publications and grant-aided research that has already been peer reviewed, we can think of these activities as *second-order* or *secondary* peer review.

Primary peer review is taken to be important because it is the principal mechanism by which scholarly quality and reliability are secured. The following two quotations are typical of claims about the importance of peer review:

3 Joan Sieber, (2005) 'Quality and value: How can we research peer review', www.nature.com/nature/peerreview/debate/op2.html. p. 2. (last accessed 9 August 2007)

4 Harriet Zuckerman and Robert K. Merton, 'Patterns of Evaluation in Science: Institutionalisation, Structure and Functions of the Referee System', *Minerva*, 9: 1, pp. 66-100 at p. 68.

5 Zuckerman and Merton, 'Patterns of Evaluation', p. 66.

“Peer review must be a central element in the scientific community’s decision-making process on scientific support.” (Royal Society, 1995)⁶

“The UK research base is widely regarded as being highly productive and by many measures is second only to the US in terms of the quality of its output. The peer review system operated by the Research Councils has underpinned this success.”⁷

Without peer review the quality and reliability of research findings cannot be guaranteed.

Peer review is important not only for the academic community but for society more generally. In recent public debates about scientific and medical studies (e.g. Measles, Mumps and Rubella (MMR) inoculations of infants; genetically modified crops; and climate change), commentators have drawn attention to the importance of ensuring that independent peer-reviewed studies form the basis for public discussion and decision on matters of science policy, rather than studies that may make newspaper headlines but not pass the test of scientific quality.

“Peer review is an essential dividing line for judging what is scientific and what is speculation and opinion.”⁸

“Given the frequent disregard in many parts of the media for accurate and balanced science and health coverage, it is not surprising that many people find it difficult to know what is factually based and what is pure junk. The fault, however, does not lie entirely with the press. Quite often journalists receive material in the form of press releases or conference papers which have not been subject to any formal vetting procedures - the critical test being that of ‘peer review’. This assesses the validity of the findings and whether the content should or should not be published.”⁹

A great deal of weight is therefore placed upon peer review in providing a sound basis for public decision-making. This weight increases when primary peer review is supplemented by secondary review as has been happening in recent years.

1.2 Primary Peer Review: the Practice

The essential principle of peer review is simple to state: it is that judgements about the worth or value of a piece of research should be made by those with demonstrated competence to make such a judgement. This principle applies to both grants and publications. With grants, an applicant submits a proposal, which is then reviewed by peers who make a judgement on its merits and eligibility for funding. With publications, an author submits a paper to a journal or a book proposal to a publisher, and peers are asked to offer a judgement as to whether it should be published. A decision is then taken, in the light of peer review, on publication or award. Box 1.1 sets out the process schematically.

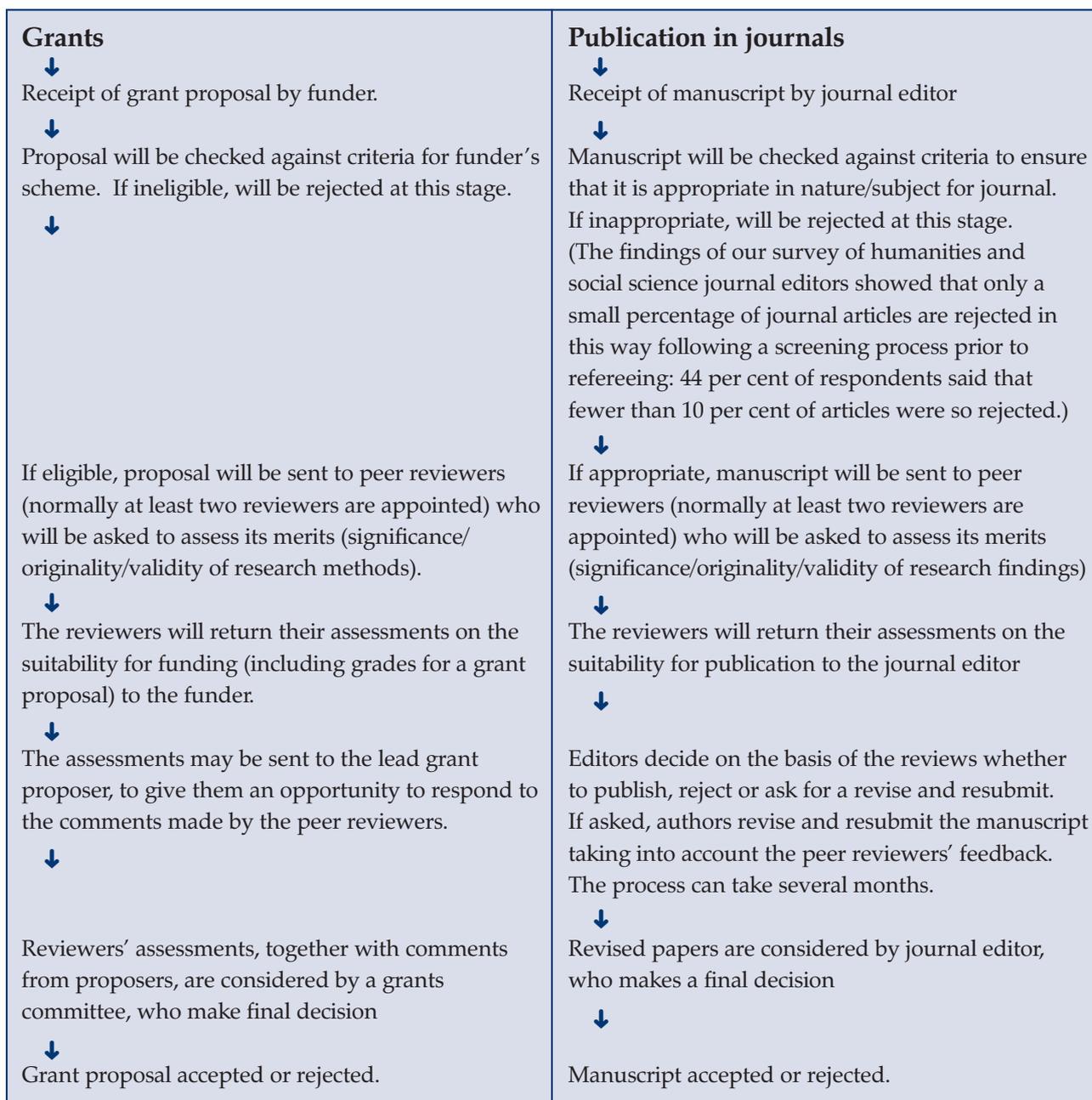
⁶ Royal Society (1995), *Peer Review - An assessment of recent developments*. p. 2.

⁷ RCUK (2006), *Project Report and Consultation on the Efficiency and Value-for-Money of Peer Review*. p. 2.

⁸ I don’t know what to believe: making sense about science stories, *Sense about Science*, 2005, p. 6.

⁹ Dr Peter Marsh, Social Issues Research Centre, as quoted on the Sense about Science web site - www.senseaboutscience.org.uk/index.php/site/project/32 (last accessed 9 August 2007).

Box 1.1: A flowchart of typical stages in the peer review process



There are of course many variants in practice on the basic principle of judgement by peers, some of which we discuss in this report. Moreover, despite the similarity between grants and publications, there are also notable differences. For example, in the case of grants there are fewer sources of awards than there are journals that publish, so that a negative decision is potentially more serious. Furthermore with grants, public or charitable money is being spent in a way that is not usually true of publications.

In any one year there are hundreds of thousands of papers published and thousands of grants awarded. It follows that academics take part in a great deal of peer review activity. Recent estimates in respect of grants across all academic disciplines suggest that the full economic costs of the peer review system is equivalent to £196 million per year (a figure calculated by combining Research Council activity with that of universities).¹⁰ The majority of the costs estimated in this figure, at c£186

¹⁰ RCUK Project Report and Consultation on the Efficiency and Value-for-Money of Peer Review, October 2006. www.rcuk.ac.uk/cmsweb/downloads/rcuk/documents/rcukprreport.pdf (page 42)

million, are those incurred by researchers and their universities when developing research grant proposals. If we add to that figure the costs of journal peer review as well as the costs of the secondary peer review, the total cost of peer review activity is considerable. If peer review is the keystone of the arch of quality, then like all systems of quality control, it comes with a price tag. For this reason alone it needs to be understood.

1.3 The Changing Context

There are a number of developments within the world of research that affect how peer review operates. Some, but not all, make the operation of the system more difficult. All are important in affecting the way in which peer review works and can be expected to work in the future.

The increase in the volume of grant applications. The volume of submissions for grants is going up. Applications for grants from the UK Research Councils, in general, have doubled since 1988. More recently, the number of research grant applications for ESRC awards has risen from 649 in 2000/01 to 1051 in 2005/06 (an increase of 62 per cent). Similarly, the number of applications for AHRC research awards has risen from 1,697 in 2002/03 to 1,942 in 2005/06 (an increase of 14 per cent). It is expected that the number of applications will continue to rise, particularly if general research funding comes to depend increasingly on grants secured. Inevitably this growth means more requirement for peer review.

The requirements for greater transparency. As competition for grants increases, so there is concern that the system of grant funding should function in a transparent way. For example, concerns have long been expressed in relation to grants that 'magic circles' should not operate. To counter such concerns, research funders need to show that their procedures provide robust evidence that good quality work is being identified and funded irrespective of its source. Use of an effective peer review system is integral to this task.

The increase in the volume of journal publications. Just as submissions to research funders are growing, so the number of journals is also growing. A survey of scholarly publishing practice in 2005 found that: "The 174 publishers analysed have launched 1,048 new journal titles in the five years to 2005, averaging 6.02 per publisher. During the same period, they discontinued 185 titles, averaging 1.06 per publisher."¹¹ Similarly, a recent report on the current state of knowledge about scholarly journal publishing in the UK said that: "There are estimated to be around 20,000-25,000 peer-reviewed scholarly journals actively being published worldwide. The volume of journals has grown consistently at a compound annual growth rate of 3-4% over the past one hundred years."¹² The increase in the number of journals published means an increase in the demand for peer review.

The development of open access journals. The development of electronic publishing means that the economics of dissemination have fundamentally changed from the print era. Dissemination is now effectively cost-free. Large research funders like the Wellcome Trust now require research publications to be made available free of charge to the public. This reflects the idea of the 'open access' journal, in which authors pay to be published (subject to peer review) rather than readers paying for copy. In a world in which there is an increasing volume of publications available without subscription, it can be argued that the quality control function of peer review has become more important. When so much is freely available, some guide is needed as to quality and value.

The growth of secondary peer review. The growth of grant applications and journal publications means that there is a large volume of work subject to primary peer review. However, more general demands

11 *Scholarly Publishing Practice: Academic journal publishers' policies and practices in online publishing*, ALPSP, 2005 (Executive Summary p. 1)

12 *Analysis of data on scholarly journals publishing*. A report commissioned in 2006 by Research Information Network (RIN), Department of Trade and Industry (DTI), and Research Councils UK (RCUK). www.rin.ac.uk/files/Executive%20Summary.pdf (p.3)

for performance assessment of institutions of higher education mean that demands on peer reviewers are augmented by the growth of secondary demands; e.g. for departmental reviews, RAE and benchmarking reviews, all of which themselves rely on peer review and tend to duplicate peer assessments already made.

Metrics rely on peer review. The Government has decided that there should be a greater use of metrics in research assessment, and HEFCE has been charged with developing a suitable system. The use of metrics might seem to relieve the demands on peer review, since metrics are aimed at lightening the load on RAE panels, the most important form of secondary review in the UK system. However, important metrics, like citation analyses, themselves depend upon primary peer review, and it is important to understand the primary system if reliable and worthwhile metrics are to be constructed.

The introduction of full economic costing for research. The introduction of FEC (full economic costs)¹³ raises questions about the peer review system. Most peer review activity is not remunerated financially, and yet the quality control function that it provides is usually thought to be essential (later, in Chapter 5, we examine some arguments in connection with this claim). How does peer review sit in a system of full economic costs? Should it and could it be properly accounted for and remunerated?

The requirements of knowledge transfer. Universities and researchers are coming under pressure from Government and funders to demonstrate that their results make a difference outside of a narrow academic context. Increasingly there is a requirement that users be involved in the funding and assessment of research. The pressure towards increased knowledge transfer has much of value, not only in the social sciences but also in the humanities. However, the requirement for non-academic involvement in the assessment of research raises important questions about who is a peer and what the role of peer users might be.

The effects of generational replacement. Although the demands on peer review and individual peer reviewers have grown in recent years, there is a process of generational replacement at work in UK universities that means that new scholars are entering academic life in a context in which peer review will play a central part in their careers, in a way that was not always true of previous generations. New entrants will be both peer reviewed and peer reviewers. How are professional norms set for these new entrants? What training and induction into the norms of peer review do new entrants receive?

The use of electronic communication. Electronic communication now means that peer review can more easily be an international process, potentially widening the range and number of reviewers. Strenuous policy efforts are being made to achieve international bench-marking of UK research, but an important force making for higher standards as judged internationally is not the imposition of a centrally determined set of goals, but the pressures that operate in a decentralised way as scholars from overseas are drawn into the appraisal of UK research and as UK researchers play a role in assessing the work of research communities abroad.

In short, peer review is operating in an academic and policy context that makes it subject to new and widening demands and in a context in which the technical and scholarly environment is changing. Some of these trends, for example the internationalisation of academic life, may be expected to strengthen peer review; other trends, for example the increasing demands upon the time of peer reviewers, can be expected to put strains on the practice.

¹³ FEC (Full economic costs) forecasts the full cost (direct and indirect costs) of undertaking a research project. Research Councils pay 80% of the full costs of the research projects that they fund.

1.4 Past Studies

Ours is not the first study to have looked at the practice of peer review. In 1989 questions were raised by the then Secretary of State for Education, the Rt Hon Kenneth Baker MP, about whether peer review was over-bureaucratic, too conservative, and too time-consuming. In response, the Advisory Board for the Research Councils (ABRC) set up a review chaired by Professor Margaret Boden, FBA. The review report (published 1990) recommended *inter alia* that peer review should be more transparent. It raised concerns that peer review disadvantaged early career track researchers, the most innovative research, and interdisciplinary research. But the review report concluded that there was no practical alternative to peer review.¹⁴

In 1995, the Royal Society published a report that focused on peer review for the purposes of awarding research council grants. It echoed the Boden report view: "Peer review must be a central element in the scientific community's decision-making process on scientific support." The Royal Society report focused on the ways in which the cost of peer review (borne largely by researchers when preparing and submitting applications) might be reduced. "As the demand from well rated proposals has increasingly outstripped the supply of funding, the peer review mechanism has become correspondingly more costly, in the sense that an increasing amount of time is consumed for every proposal ultimately funded, and more frustrating and demoralising for all involved." The report argued that peer review had to "operate at an acceptable efficiency" and considered proposals on ways in which the volume of demand might be moderated e.g. restricting the number of applications from each institution; allocating funds through larger blocks of grants; and pre-screening outline proposals.¹⁵

1.5 Principal Questions

In conducting our enquiry, our first concern was to gain an idea of how peer review is functioning at present. How does it work in decisions on grant making or publication? What difficulties are those operating peer review finding? What opportunities arise? (Details of our methods of working can be found in Appendix. 1.)

Although peer review is central to the scholarly enterprise, it has attracted various criticisms, in addition to the concerns of cost that prompted the earlier enquiries to which we have referred. At various points we consider some of those criticisms. They have some merit, but we echo Sieber's view, cited at the head of this chapter, that they imply the need for improvements in practice rather than abandonment. We have therefore been concerned to identify the principles of good practice and to examine the extent to which there are models of good practice that are worth disseminating. However, as we worked on the report, we became aware of how many issues in research policy (in particular research training, full economic costing, metrics and concerns about innovation) bear upon and are affected by peer review as a practice. It is these concerns that we have also sought to address in this report.

In preparing this report, we have become aware of how far peer review is discussed in very general terms across all academic disciplines or is discussed with an implicit model drawn from the practices and processes of medicine and the natural sciences. We accept that there are general issues, and we also acknowledge that, in many respects, much can be learned from the study of peer review in medicine and the natural sciences. However, there are special issues that are raised for the humanities and social sciences, and it is these to which we have tried to give special attention. We are also

¹⁴ *Peer Review: A report to the Advisory Board for the Research Councils from the Working Group on Peer Review*, ABRC, 1990, p. 1.

¹⁵ *Peer Review - An assessment of recent developments*, Royal Society, 1995, pp. 2-3.

conscious that there is very little information that is specific to the humanities and social sciences, and urge those who commission large-scale studies across all academic disciplines, to look at ways in which the resulting data can be disaggregated by discipline.

Publication practices in many areas of the humanities and social sciences differ from those in the natural sciences. The volume of journal articles produced tends not to be as high as that for the sciences, the production cycle is much slower, and, for many subjects, the monograph remains a more highly valued output. Scholarship in many fields within the humanities and social sciences is less driven by the latest published articles: work published 30 to 50 years earlier can often still have resonance. It often takes longer to assess the contribution made in the humanities and social sciences, particularly by a major monograph. Book reviews can play a central role, followed by a process of digestion, reflection and debate. The growing authority of a major work takes shape in the treatment it is given by other authors. This can take five to ten years. Often it is in book form - a major appreciation or alternative world view, sometimes sparked off by the previous author. In fast moving empirical science, where journals are the key medium (Economics included), this is not as true. Not only do the humanities and the social sciences differ from medicine and the natural sciences, they are also diverse with respect to one another.

By contrast with the natural sciences or medicine, where English is the *lingua franca*, there are a number of fields within the humanities and social sciences in which the principal medium for publication is in a language other than English. Some of these areas are highly specialist, and one would expect the authority of particular pieces of work to be established by means other than the numerical volume of citations.

With regard to research grants, much high-quality work in the humanities and social sciences is delivered by researchers who do not use expensive equipment or work in large-scale teams. Research is often undertaken by lone scholars, whose needs may be relatively modest - chiefly time freed from other duties to pursue their research. As a consequence, small grants remain important for these disciplines (as the British Academy has repeatedly found with the academic support for its own small grants programme), and so the overhead costs for peer review may be high relative to the size of the award. There are also often small pockets of research excellence in the humanities and social sciences, as there are not the same concentrations that frequently occur in medicine and the natural sciences.

With these distinctive features in mind, we have tried to set forth an understanding of peer review, particularly as it is relevant to the humanities and social sciences, in this report. We begin with what we have termed primary peer review in relation to publications and grant awards, topics that are covered in the next two chapters. We then go on to reflect upon the significance of peer review in relation to training, full economic costs, metrics and innovation.

Chapter 2: Peer Review for Publication

“Peer review does not, and cannot, ensure perfection: scientific journals are records of work done and not of revealed truth. If they were to insist on absurdly high standards science would suffer more than it would gain, purchasing reliability at the expense of innovative quality.”¹⁶

“Whenever surveys of scholars’ opinions about scholarly publishing have taken place, the maintenance of peer review systems has been a top priority requirement (The Royal Society, 1981; Rowland, 1982; McKnight and Price, 1999; ALPSP, 1999, 2001, 2002). The graded quality control system provided by a variety of journals in a “pecking order” is helpful to readers by enabling them to identify the most important papers, whereas an undifferentiated mass of publications on the Internet would overwhelm readers by its volume. And finally, in addition to their “accept/reject” judgement, referees improve papers before publication.”¹⁷

2.1 Introduction

One of the principal uses of peer review is to enable journal editors and publishers to make decisions on which articles or monographs are to be published. As an editor of monographs said to us in submitted evidence: “Peer review is completely fundamental to our activities. It is the major guarantor of quality control that we have...” Similar sentiments were repeatedly echoed to us by journal editors.

Although peer review is used to assess both journal articles and scholarly monographs, there are differences of practice between these two forms of publication. With articles, the review is usually of work that is completed - or at least presented as complete, since peer review itself may lead to requests to rewrite and resubmit. In the case of monographs, publishers will often make decisions based on sample chapters and the post-publication book review has provided one of the means by which quality has been assessed.

Across the humanities and social sciences, there are considerable differences in the importance attached to the journal article as distinct from the monograph. In some subjects, such as Economics and Psychology, the journal article has now become the principal form of publication. In many other subjects, by contrast, the monograph remains the touchstone of quality. In this chapter, we focus on peer review in relation to journal articles. This is not intended to presume a view about the relative merits of the different forms of publication. That is a matter of evolving convention within each discipline. Rather our focus is on journals because this is the area in which there have probably been most developments of practice in recent years, both in the volume of publication and in the way that journals operate. We are conscious of the importance of monographs, however, and in a later chapter we pay particular attention to the need to ensure that they receive their full weight in research assessment.

2.2 What is Journal Peer Review?

We have seen in Chapter 1 that the essence of peer review is straightforward. In the case of journals, it is the practice by which papers submitted for publication are reviewed by peers, that is to say those with recognised expertise in a field, who are asked to recommend to a journal editor whether or not the paper should be published. In this general sense, peer review is a common practice not

16 Lock, S. (1985). *A Difficult Balance. Editorial peer review in medicine*. ISI Press, Philadelphia.

17 F Rowland, *The Peer Review Process: a report to the JISC Scholarly Communications Group*, p.11.

only across the humanities and social sciences but also across subjects in medicine and the natural sciences. A recent report defines peer review as “the evaluation of scientific research findings or proposals for competence, significance and originality, by qualified experts who research and submit work for publication in the same field (peers).”¹⁸ With the relevant modification, this definition easily fits the humanities and social sciences.

However, once we move beyond this initial characterisation, there are many different ways in which this very general principle can be implemented, two features of which are particularly important.

- (1) There is often not a sharp separation between the role of editor and the role of reviewer. Where reviewers disagree, for example, it may fall to the editor to make a judgement in the light of conflicting advice. Similarly, many journals practise an initial stage of editorial review in which an editor will decide whether a submission conforms to the journal’s editorial policy.
- (2) The meaning of peer review can be extended to include not only pre-publication decision-making, but also post-publication assessment. For example, in the social sciences in recent years there has been a great deal of interest in the extent to which reported statistical findings can be replicated by other researchers.¹⁹ With the growing availability of electronic databases, original data can now be made available to other researchers to provide a check and assessment of published findings. Similarly, review articles surveying a particular sub-field may well make a contribution to the assessment of a particular piece of work.

Aside from these two particular points, peer review is practised in many different ways by journals in humanities and social sciences. Differences cover such matters as pre-screening before sending to referees, participation by editors themselves in the process of assessment, who makes the final decision and whether authors remain anonymous. Box 2.1 describes a range of practices outlined by journal editors who responded to our survey.

Anonymisation is a particularly interesting example of how practice may vary depending upon the conventions of different disciplines and the policies of different journals. A number of journals in recent years have moved towards the practice of not only protecting the identity of the referees from authors but also of preventing referees from knowing the names of authors (so called ‘double blind’ reviewing). However, not all journals or disciplines are equally committed to peer review in this form. In disciplines like Economics, for example, where it is common for researchers to place working papers on personal web-sites, it may be impossible to keep the identity of an author anonymous from the peer reviewers.

Another complication in the humanities and the social sciences is concerned with the definition of who is a peer. In the sciences and in many subjects within the humanities and social sciences the test of who is a peer will be clear: a peer is an academic of relevant standing in the discipline. However, in fields where practitioner or policy relevance is important, a suitable reviewer may be someone from a non-academic background with relevant experience.

¹⁸ Sense about Science, 2004.

¹⁹ See, for example, Gary King, ‘Replication, Replication’, *PS: Political Science and Politics*, 28:3 (1995) pp. 444-52.

Box 2.1: Practices of Journal Peer Review (drawn from the responses to our survey of HSS journal editors)

Pre-screening process prior to refereeing	While 90 per cent of the humanities and social science journal editors who responded to survey rejected articles in a pre-screening process, the proportion rejected was low: 74 per cent of the respondents using some form of pre-screening process between 1 and 20 per cent of articles in this way.
Participation by editors in the peer review assessment	Only 7 per cent of respondents never participated in the refereeing process. Editors would participate: (a) if article fell within their field (44 per cent of respondents); (b) if casting vote was needed (12 per cent); or (c) in difficult cases (37 per cent).
The number of referees appointed	It was unusual for only one referee to be appointed: 11 per cent of respondents followed this practice, several of whom did so in addition to the editor who would also provide an assessment. While the number of referees may vary, the number most frequently given by our respondents was (at 70 per cent) two.
Anonymity of authors	61 per cent of respondents said that papers were always refereed with the author being anonymous
Anonymity of referees	89 per cent of respondents said that the referees were always anonymous, and 94 per cent of respondents believed that the identity of referees should not be known by the author
The selection of referees	The answers we received suggested that the system is very dependent upon an editor's personal networking skills and knowledge of the field.
The use of overseas referees	Overseas referees appear to be frequently called upon to assess work: 56 per cent of the respondents said that they either 'always' or 'usually' sought assessments from overseas.
Conflicts of interest	The responses we received suggested that humanities and social sciences editors rely heavily on referees to volunteer information about whether there might be a conflict of interest. Only 6 per cent of respondents required referees to complete a declaration regarding conflict of interest.
Who makes the final decision on acceptance or rejection	The editor(s)/editorial board make the final decision according to 99 per cent of our respondents.

2.3 Is Peer Review Needed?

There are some criticisms of peer review, with some so fundamental that they call into question the practice itself.

Peer review slows up the publication process. One submission to our review suggested that the original purpose of peer review was to ration access to resources for scholarly exposure. Nowadays, however, exposure is not a scarce resource, since publications can be made available electronically, essentially free of cost. The question therefore is one of quality control and we do not know how much refereeing the scholarly market actually wants.

“Scholarly scrutiny depends on a great deal of selfless work by referees and editors. On the whole, it works ‘as advertised’ and its objective and ethical standards (in my experience) are satisfactory or better most of the time. But it also consumes a great deal of time, and slows down dissemination. Furthermore, such great reliance on goodwill is costly, perhaps inequitable, and perhaps also wasteful... there is a question whether the rising opportunity cost of gatekeeping remains justified, whether other reputational mechanisms might not be cheaper, and whether readers should not be allowed more leeway to decide for themselves what is and isn’t good enough - as is more the case in the book market.” (comments made by a humanities and social science researcher)

Peer review allows for irresponsible behaviour on the part of the referees. Many academics have a horror story about journal peer review: the paper had obviously not been read; the referee used the occasion to promote their own work; wounding or malicious statements were made, some of which might damage, say, new entrants to academia.

Peer review inhibits innovation in published work. Peer reviewers are part of the gate-keeping of published work. As such peer review can be a vehicle for pressures towards conformity to established research, which may drive out innovative work.

These points are well taken as criticisms of the practice and as indicators to the way in which it might be improved, but do they amount to a case for rejecting peer review as a practice? In our view such arguments provide reasons for being concerned with the efficient and effective practice of peer review, but not for rejecting it as a matter of principle for the following reasons:

- (1) Although peer review does slow up the publication of papers, that is an unavoidable consequence of the system of quality control. Since papers can be made available electronically on pre-publication web-sites, they are accessible to those for whom the quality certification is less important. This can happen for quite legitimate reasons, say because the empirical results, which are uncontested, are more important than their theoretical implication, which may be the point at issue for the referees.
- (2) The peer review process should not be seen simply as a selection mechanism, which it is, but also as a way of improving the quality of the paper. This was a view put strongly to us in our meeting with journal editors.

“Authors can, and do, express gratitude for this service. Of course, they would be even more grateful for the help in improving quality and a fast turn-around time, but that comes back to the issue of practicality.” (humanities and social science journal editor)

This view was widely echoed among our respondents.

- (3) The movement to make all journals open access, so that the material they contain is freely available to anyone who wishes to examine it, might suggest that the role of peer review should be reduced, since readers would incur no economic costs in making up their own minds about the value of published work.

However, as one commentator has put it, there is “even in the electronic future, a place for a repository for scientific papers of certified quality that has a long lifetime, stable contents, and a known means of access.”²⁰ In short a future that looks like the existing peer review journal. Indeed, it is possible to argue that in an electronic world, peer review becomes more, not less important, in its role of guaranteeing quality.

For these reasons, we consider that peer review remains an essential, if imperfect, practice for the humanities and social sciences.

2.4 Principles of Good Practice

However, to accept the value of peer review does not mean that it works well in practice. In this section we consider what principles are implied by a commitment to peer review and we consider some problems that can call into question the integrity of the practice.

We have already noted that humanities and social science journals adopt various approaches. We see these models as reflecting a diversity of subject matter and approaches in these subjects, and we would not suggest that there was one model that all should follow. Indeed, *it is a considerable merit of the way in which the peer review system works that there is not one single model of good practice that all should follow, but instead decentralised diversity.*

However, and despite the desirable diversity, it is possible to state principles general enough to apply to a variety of practices, and which, if followed scrupulously, would make a difference to the overall quality of academic publication. The three principles in particular that we consider are: timeliness, transparency; and verifiability.

Timeliness. Timely publication is important for a number of reasons. The swift and effective dissemination of research findings is at the heart of the academic enterprise. Increasingly, in a context in which demonstrable performance is a crucial condition of academic appointment and promotion, entrants to the academic profession depend upon timely judgement. The increase in the range and number of journals is also relevant. At one time journals in a field were, in effect, monopoly or near-monopoly gatekeepers of publication, since there were few other outlets. In many humanities and social science subjects, journals have multiplied, so that a rejection from one journal means that authors can turn to others. However, that in turn implies that authors in fairness ought to know as quickly as possible if their paper is to be rejected, in order to provide them with the opportunity to take their work elsewhere.

The difficulty is that quick decisions cannot be guaranteed where peer review is involved. In part this is due to the vagaries of referees and journal offices. Where specialist judgement is needed, there may only be a few people in the world who are able to offer a peer judgement. For perfectly valid reasons, they may not be readily available. Even where a referee is available, the amount of effort involved may be considerable. In this regard, some branches of the humanities and social sciences are closer to the world of Mathematics than they are to other science subjects. In Mathematics, peer judgement is needed as to whether a proof is valid or not. Similarly, in the humanities and social sciences a judgement may be needed as to whether a conceptual or logical argument is valid or really adds value, and verification of this may take some time.

²⁰ Thomas von Foerster, ‘The Future of Peer Review’, American Academy of Arts and Sciences, 2001, available from <http://www.amacad.org/publications/trans5.aspx>

None of these considerations means that timeliness is unimportant or that unnecessary delays can be excused. In particular, referees need to understand the importance of timeliness and editorial offices need to ensure that timely delivery of reports is an element of choosing referees. Authors are increasingly willing to pressure editors for a decision within a reasonable amount of time, which in many cases is perfectly reasonable.

Transparency. Transparency is important because it is a central element of the scholarly enterprise. The possibility of criticising or refuting established ideas and approaches is at the heart of creative scholarly work. One of the advantages of the peer review system is that referees provide reasons why an article should be published or not. Editors have a responsibility to indicate the basis upon which they have made their decisions. Sometimes these decisions will not be pure quality decisions, but judgements about the extent to which the work under consideration is consistent with the editorial policy or remit of the journal.

The principle of transparency does not necessarily imply a particular range of practices. Some commentators on the peer review system have argued, for example, that the whole system should be totally transparent, with referees knowing the name of authors and authors knowing the names of referees.²¹ (Interestingly, there was some support for such a regime among the postdoctoral fellows to whom we talked.) However, we think that there are advantages in maintaining the practice of anonymous referees (allowing those referees to waive their anonymity if they so choose), if candid reports are to be written. This is likely to be particularly important in sub-disciplines where there are only a small number of practitioners.

Equally, we do not see transparency as requiring editors merely to be post offices for referees' reports. Some reports can contain malicious or wounding statements that are at best incidental to an academic judgement of the quality of work being assessed. Such statements can be particularly damaging to new entrants. Editors should not feel under an obligation from the principle of transparency to pass on such reports without amendment. In some cases, even, they may choose to draw a veil over the report completely - even if some element within it is academically important.

Verifiability. Our third principle is that of verifiability. As we noted in the opening section of this chapter, peer review in an extended sense involves post-publication evaluation of work as well as pre-publication evaluation. A good system of peer review will facilitate post-publication evaluation. Among other things, this means that articles should conform to good citation practices, so that readers can locate the contribution the particular article makes to the wider scholarly discussion. It should also involve access to replication data sets, so that others can re-run and evaluate the findings.

An example of the importance of verifiability is shown in Box 2.2.

21 Richard Smith, 'Opening Up BMJ Peer Review', *British Medical Journal* 318 (7175), pp. 4-5.

Box 2.2 Abortion and the Drop in Crime Rates

In the May 2001 issue of *The Quarterly Journal of Economics*, John J. Donohue III and Steven D. Levitt published an article discussing the effects of legalised abortion on crime rates in the United States. The core of their argument was that the legalisation of abortion in the wake of *Roe vs. Wade* had led to the marked reduction in crime rates from 1991 onwards. They concluded by saying that their estimates of the effects meant “that legalized abortion is a primary explanation for the large drops in murder, property crime, and violent crime that our nation has experienced over the last decade.” They went on to say that legalised abortion “may account for as much as one-half of the overall crime reduction.”²²

Donohue and Levitt posited two mechanisms by which legalised abortion may have been responsible for the drop in crime rates. Firstly, since crime is mainly committed by young adults, a reduction in the numbers in a given cohort will reduce the numbers in the population at large who are likely to commit a crime. Secondly, foetuses aborted were likely to have been unwanted, and unwanted children facing disadvantage are more likely to commit crimes, so that the *selective* effects of abortion on the cohort reinforce the absolute numbers effect.

In 2005 two economists at the Federal Reserve Bank of Boston, Christopher L. Foote and Christopher F. Goetz, published an electronic paper pointing out some problems with part of the statistical analysis that underlay the claims about legalised abortion and crime rates.²³ In particular, they pointed out that the reported results did not control for the effects of the year and the state in which the results were recorded. This was important, because there may have been circumstances specific to particular states in particular years that affected crime rates. For example, the use of crack cocaine spread across different states in different years in the 1980s, and unless one controlled for these circumstances, no reliable inferences could be drawn. Foote and Goetz went on to show that if one did perform the necessary statistical controls, then the posited effect between legalised abortion and crime rates disappeared.

In a reply to the critique, Donohue and Levitt acknowledged the error.²⁴ They argued, however, that on a measure of abortion rates superior to the one they originally used (one which allowed for cross-state mobility in access to abortion services) the original hypothesis of a relationship between legalised abortion and reduced crime could be upheld.

At the time of writing, a version of the Foote and Goetz critique is scheduled to appear in *The Quarterly Journal of Economics*.

What are the lessons for our understanding of peer review?

- Pre-publication peer review will not pick up all the problems with a paper. The failure to run the tests on state/year effects was only spotted after publication, despite the journal in question being of high standing.
- Post-publication review worked. Not only did a reader spot the error, but the critics were able to gain access to the data and computer programs from the original authors to cross-check the results. This is an example of good practice.

22 John J. Donohue III and Steven D. Levitt, ‘The Impact of Legalized Abortion on Crime’, *The Quarterly Journal of Economics*, CXVI: 2 (2001), pp. 379-420 at p. 414.

23 Christopher L. Foote and Christopher F. Goetz, ‘Testing Economic Hypotheses with State-Level Data: A Comment on Donohue and Levitt (2001)’, Federal Reserve Bank of Boston, *Working Papers No. 05-15*, available at <http://www.bos.frb.org/economic/wp/index.htm>.

24 John J. Donohue III and Steven D. Levitt, ‘Measurement Error, Legalized Abortion, the Decline in Crime: A Response to Foote and Goetz (2005)’, mimeo, January 2006.

2.5 Conflicts of Interest

Many conflicts of interest arise more readily in medicine and the natural sciences than in the humanities and social sciences. Examples in medicine, for instance, arise from research financed by pharmaceutical companies. However, within the humanities and social sciences, even if journals follow principles of good practice in peer review, there are bound to be problems that arise in the conduct of peer review assessment. One important sub-set of such problems stems from potential conflicts of interest between authors and referees.

One frequently cited issue in this context is the extent to which there is a conflict of interest between a referee and an author if the referee has seen a version of the same paper (or indeed exactly the same paper) before, when acting as a referee for another journal. There are some who feel that in these circumstances, it is the duty of the referee not to assess the paper, whereas others take a more liberal view. One difficulty with saying that the same person should not act as referee twice on the same paper is that the referee will typically explain to the editor why he or she cannot act, and this explanation will make reference to the fact that they have seen the paper before, together with an indication of the view that was taken (and it may be of course that the view was favourable to publication). Editors in this situation may well take the view that they would like someone else to look at the paper, but they may also take the view that they would like to see the original report. After all, if the original report made valid criticisms of an article, and the article has not changed, then an editor may feel that the criticisms are a good reason to reject.

Another type of conflict of interest arises when author and referee are working on similar material, and the referee gains valuable information or understanding by virtue of the capacity in which he or she acts. In such cases, referees should inform the editor that they are unable to make a fair judgement about a paper because they are working in such a closely related area that they fear their judgement would be compromised by fear of pre-emption.

Authors too may have reasons to worry about securing their scholarly rights in a piece of work. Many studies in the humanities and social sciences depend upon a great deal of data gathering, and data sets are held electronically. In principle such data sets provide a rich resource for researchers, since they can be widely shared. However, those who have collected the data may feel the need to protect their intellectual investment in that data to the point where they have published what they can from it. There could be problems, for example, if a referee asks to see a data set to perform some sensitivity tests as part of the refereeing process.

Given the diversity of the disciplines within the humanities and social sciences, it is difficult to formulate a simple code of practice dealing with conflicts of interest that can apply to all journals. We endorse the Blackwell view that “editors should adopt a policy about conflicts of interest that best suits their particular publishing environment, and should describe this in their editorial policy.”²⁵ We also urge Professional Associations to ensure that this issue is addressed in their codes of conduct.

2.6 Scholarly Malpractice

There is no reason to think that scholarly malpractice is widespread, but it is serious when it occurs and has effects out of proportion to its prevalence. Malpractice can take a variety of forms. One is the fraudulent presentation of results that were not obtained. In medicine and the natural sciences this form of malpractice has led to the shaming of those who had previously enjoyed

²⁵ *Best Practice Guidelines on Publication Ethics: a Publisher's Perspective*, Blackwell Publishing Ltd, 2006, (Section 8.11)

high prestige, including Jan Hendrik Schon, who had been tipped to be a Nobel Prize winner²⁶ and Hwang Woo-suk.²⁷

Another problem is plagiarism and other derivative text, the presentation of other peoples' work as though it were one's own, sometimes extending to whole papers. In recent years several journals have had to withdraw papers because they had been plagiarised in whole or in part. Lesser, though still important, forms of malpractice include 'redundant' publication, where the same material is presented by an author in more than one publication and 'salami slicing', by which research findings are stretched out over more publications than are necessary.

Peer review is sometimes criticised because examples of scholarly malpractice slip through the net. We do not share this view. It is not realistic to expect peer reviewers or editors to identify all cases of fraud or deception. Nor is it their responsibility. Responsibility for avoiding plagiarism and redundant publication rests with authors themselves. We note, however, that, though journals are generally good at reporting individual instances of malpractice, mechanisms for reporting instances of malpractice more generally are ad hoc (usually in the form of emails to other editors, giving details). There may be a role here for professional associations or HE institutions to facilitate the repositories of information regarding instances of malpractice thus strengthening the peer review system.

In this context, software developments may help, enabling editorial teams to run submitted articles through a checking system. But no software can check a submitted article against an unpublished source. The peer review system necessarily rests upon trust, raising questions of motives and incentives (see Chapter 5).

However, we need to remember that judgements of individual malpractice involve claims about complex circumstances and if substantiated have severe implications for the individual involved. For this reason, we concur with the view of Hames as noted in Box 2.3.²⁸

Box 2.3 Dealing with Plagiarism

"Editors and journals must, however, always be very careful to distinguish between genuine errors and the intention to deceive; the latter constitutes misconduct, the former does not. They should not make or spread allegations of misconduct before investigating those allegations and finding them to be substantiated. Not only can this seriously, and perhaps unjustly, damage a researcher's reputation, it also opens up the possibility of litigation. It must be recognised that some apparent 'misbehaviour' may be the result of ignorance of good practice. This can happen, for example, with authors who do not have much experience of research or publishing, or with very junior authors who have not received much, or perhaps any, guidance from their supervisors in these areas."

26 A. Mulligan, 'Is Peer Review in Crisis?', *Perspectives in Publishing*, 2, pp. 2-6, at p. 2. (see page 47 for full web page URL).

27 Anna Fifield and Clive Cookson, 'Seoul Searching: Koreans find their Rapid Development has Hard Scientific Limits', *Financial Times*, Thursday January 19 2006, p. 15.

28 Irene Hames, 2007, *Peer Review and Manuscript Management in Scientific Journals: Guidelines for Good Practice*, Blackwell Publishing Ltd, p.173.

2.7 Conclusions

- (1) Journal peer review is not a single practice but a set of practices. All these practices can be judged by how well they serve the two principal functions of peer review.
- (2) Those two functions are: quality control and enhancement. The first operates by selection; the second through constructive criticism.
- (3) While there are legitimate criticisms about the way in which peer review currently operates in respect of journals, there is no alternative process by which quality judgements can be made in selecting articles for publication.
- (4) Diversity of practice is a strength, not a weakness.
- (5) While there is no one model of good practice that journals ought to follow in respect of peer review, there are general principles to which practices should aspire, and these include timeliness, transparency and verifiability.
- (6) As pressures to publish increase, journals will face increases in potential conflicts of interest. Again, there is no one model to follow (say in respect of the second refereeing of an article that a reviewer has already seen), but journals in the humanities and social sciences should develop policies to deal with conflicts of interest, and should be explicit about what their policy is. Professional associations may have a role to encourage informed reflections on what should be regarded as a conflict of interest in their discipline.
- (7) While peer review is one way by which scholarly malpractice can be identified, it cannot be a guarantee against it. Professional associations or HE institutions may also have a role in facilitating the development of repositories on such instances of scholarly misconduct.

Chapter 3: Peer Review and Grants

3.1 Introduction

The peer review processes used to assess research grant applications are similar in many respects to those for publication. The basic principle, namely that a judgement should be made by those with demonstrated competence, remains the same in both cases. However, there are four key important differences of practice:

- (1) Grant decisions are based on judgements of the prospective quality of the research that will result from an award (if successful), rather than retrospective judgements on the success of research undertaken.
- (2) By contrast with those journals that practise double-blind reviewing, the identity of the applicant is not hidden from the peer reviewers, since track-record is often, and legitimately, used as one of the indicators of potential future success.
- (3) By contrast with journal submissions, there are fewer alternatives to an unsuccessful grant application.
- (4) Public money is involved in the award of grants, either directly or through the tax expenditures from which research charities benefit.

These differences do not affect the central value of peer review, namely the making of judgements on the basis of the intellectual and academic quality of the work, but they do raise points where issues of principle arise in the way peer review is practised.

3.2 AHRC and ESRC Peer Review Processes

Within the humanities and the social sciences, the AHRC and the ESRC are the principal funders, and during the course of our enquiry we spoke with officials and academic members from both bodies. We were particularly concerned to ascertain what they saw as the strengths and weaknesses of the systems they operated as well as how they responded to criticisms that have been levelled at the way the Research Councils operate.

Box 3.1 sets out the flow chart of judgement for large grants in the AHRC, and Box 3.2 the flow chart for the ESRC standard grants.

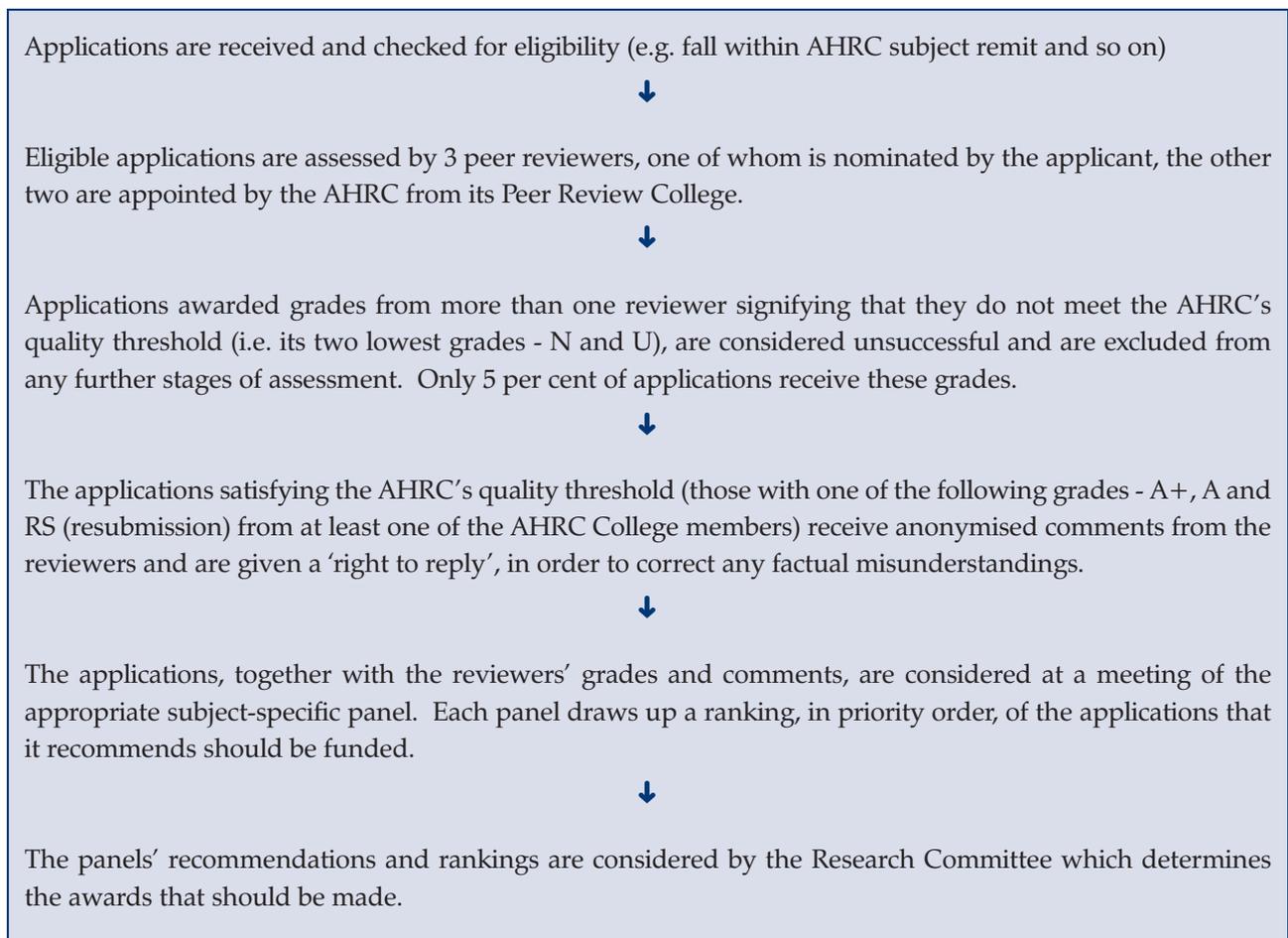
The first three stages in Box 3.1 also apply to the AHRC's small grants scheme (applications seeking grants less than £20K), but there is no 'right of reply' stage and the grades and comments of applications with an 'agreed' alpha grade of some kind are referred to the convenor of the appropriate subject-specific panel, rather than to the full panel. The convenor determines the final grade to be awarded, and the application's ranking. In the case of interdisciplinary applications, grades and comments are referred to more than one convenor. The Chair of Research Committee considers the convenor's grades and rankings and confirms the final funding outcome.

Similarly, the ESRC operates a lighter touch assessment of the applications to its small grant scheme, defined as any request up to £100k. These applications are assessed by a member of the ESRC's Research Grants Board, and also a member of its Virtual Peer Review College. Their grades and comments are then forwarded to the Chair or Vice-Chairs of the ESRC's Research Grants Board who determines the final funding outcome.

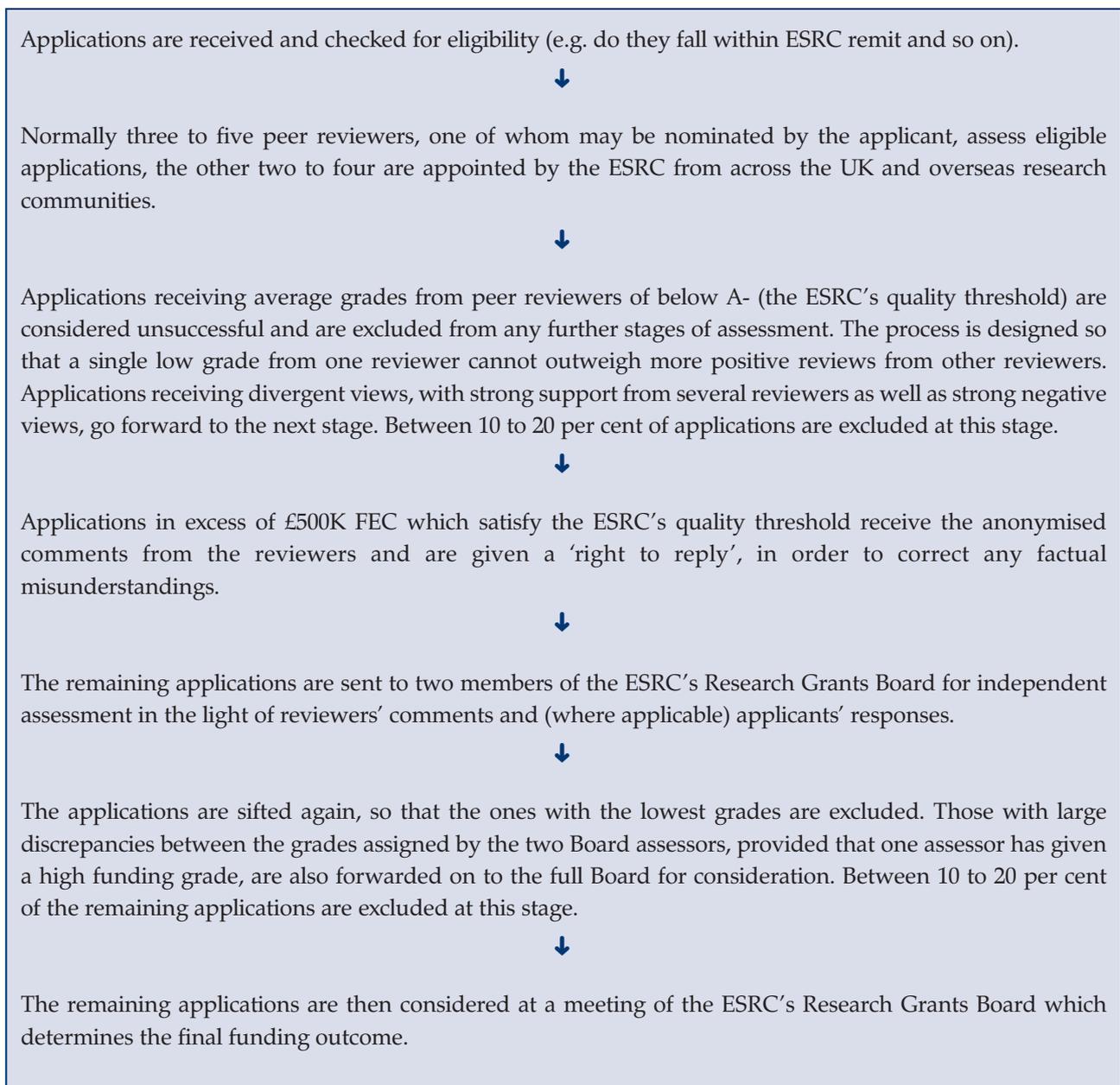
3.3 British Academy Peer Review Processes

The British Academy itself is a research funder, and makes awards in responsive mode to all branches of the humanities and social sciences. As with the ESRC and the AHRC, the Academy makes sure that the number of peer reviewers is proportionate to the size of the award. Research grant proposals are assessed by subject specialists (mostly Academy Fellows) in the light of references and any external evaluations. Assessors' comments and recommendations are forwarded to a final award committee, which considers the applications and decides on awards to be made in the light of the quality of the competition and the budget available. There may be an interim stage where applications are reviewed by a selection panel. For larger research grants, feedback is given to unsuccessful applicants from the external evaluators (where provided). For other schemes, feedback is provided only if the assessors or the final award committee have specific comments/constructive advice for the applicant. There is currently no 'right to reply'. Academy Fellows undertake this service for free - there is no charge on the public purse. As Fellows are elected on the basis of their academic distinction, the Academy can be confident that its peer reviewers are leading researchers in the field.

Box 3.1: A flowchart of AHRC peer review processes for large grants



Box 3.2: A flowchart of ESRC peer review processes for standard grants (more than 100K)



3.4 How is Peer Review Regarded by the Research Community?

Support for peer review remains high within the research community. "93 per cent of university researchers believe that the peer review system is worthwhile despite the amount of effort involved."²⁹ While recognising that peer review is not infallible, the majority of the research community draw attention to the central part that it plays in academic life and hold peer review in high esteem. In our discussions with AHRC panellists, a key point raised was that this enquiry should endorse the principle of peer review.

Despite the generally high esteem in which the peer review process is held, we are aware of concerns that it is difficult to find suitably qualified peer reviewers to serve on either the AHRC or the ESRC's

²⁹ Finding of an extensive survey of academic community conducted in 2006 on behalf of Research Councils UK (RCUK)

Colleges for peer reviewers. These concerns were supported by the ESRC reviewers we met, who did not believe that the most talented researchers always engaged in peer review activities for the Research Councils, because they are under pressure from their host institutions to focus on publishing their own research. We seek to address some of these issues in Chapter 5, where we discuss incentives and the need to ensure that peer reviewing as an activity is accorded the status that it deserves.

We are also aware that the difficulty of recruiting suitably senior scholars for regular peer review activities can lead to a situation in which the legitimacy of grant awards depends upon the representativeness of the reviewers. While considerations of subject, institutional, geographical, age, and gender balance are significant, these factors should not override the prime criterion for the selection of peer reviewers - academic distinction.

One reason this point is especially important in the case of grants is that there is no alternative to the peer review process that does not have severe disadvantages. Without peer review of individual applications, funders would have to rely upon track record, an approach that would work against the development of new entrants to the research community. The point has been made well by Christophe Schneider from the Deutsche Forschungsgemeinschaft:

“Peers are human. Consequently, peer review has been faulted with all the epithets applied to human nature by disgruntled humans: Unreliable, irreproducible, biased through self-similarity or sheer greed, sexist, ignorant, careless, dishonest. So peer review easily appears as the worst form of underpinning decisions on distributing scarce resources to the best research - until alternatives are seriously considered. Consider, for instance, political correctness, formal bureaucracy, first come first served, lottery systems, citation analysis, journal impact factors, or enlightened arbitrariness. In the author’s view therefore, peer review is the best of all known methods of underpinning decisions on competing scientific proposals as long as it is properly managed.”³⁰(Christoph Schneider, Deutsche Forschungsgemeinschaft - DFG - October 2006).

This is a view that we endorse.

3.5 Examples of Good Practice in Quality Assessment

As the Working Group engaged in discussions with those carrying out peer review for the Research Councils, it was impressed with examples of good practice. Three are worth mentioning here.

The AHRC has eight subject-specific panels, which meet on a regular basis to consider research grant proposals, together with the grades and comments that have been assigned by the peer reviewers appointed to assess proposals. Each panel determines the final grade to be awarded to proposals and ranks each one in order of priority. The panels’ recommendations and rankings are then considered by the AHRC’s Research Committee which determines the final funding outcomes.

In discussion with a group of panel chairs, the Working Group was impressed by certain aspects of this system. The panel meetings enable the panels to reach independent judgements of the quality of the proposals under consideration, allowing them to consider each proposal on a case-by-case basis, together with the whole range of comments and grades that have been assigned. Under this arrangement, peer reviewers’ judgements are treated seriously but not mechanistically: the

30 Christoph Schneider, Deutsche Forschungsgemeinschaft - DFG - October 2006).

AHRC is able to avoid a formulaic approach of averaging grades, so one low mark should not automatically rule out a proposal. The panellists the Group interviewed were confident that the process meant that they were able to identify the reviewers whose grades tended to be inconsistent (either too high or too low). While this kind of structure involves extra costs, the additional benefits that it brings far outweigh these considerations.

The second example is the way in which digital resources are assessed and evaluated. At present, the peer review systems for e-resources are in their infancy. Given the limited funding available, it is essential that peer review systems should be able to identify with precision the resources that are worth creating and sustaining and those that are not. A recent report arising from a study funded by the AHRC, 'Peer review and evaluation of digital resources for the arts and humanities', recommended the establishment of a framework for evaluating the quality, sustainability and impact over time of digital resources for the arts and humanities.³¹ The project, led by Professor David Bates and Dr Janet Winters, Institute of Historical Research, recommended *inter alia* that there should be a consistently applied system of peer review of both the intellectual content and technical infrastructure, as well as on-going evaluation through the whole life of a project at different stages in its cycle. It was claimed that the 'value-added' of this whole cycle evaluation process outweighs the increase in peer review costs that would be incurred. The recognition that as disciplines evolved, technical and academic judgements needed to be brought together in ways that has not previously occurred was especially important.

The third example is the ESRC's practice (once final award decisions have been reached) of sending reviewers a copy of all the other reviewers' comments on that proposal (in anonymised form), together with a letter informing them of the final funding outcome. This initiative has a number of objectives: improving the transparency of decision-making; improving feedback to reviewers; and helping reviewers to develop their peer-reviewing skills.

3.6 Transparency of Peer Review Practices

As with journals, the principle of transparency is important for research funding, indeed more so given the limited number of funders in comparison with journals. This issue was considered carefully by the Boden report in 1990, which recommended that Research Council peer review practices should be more transparent, and highlighted the importance of feedback in order fully to utilise the academic time spent on peer review.

"Councils should not underestimate the degree to which obscurities in their systems can have a depressing effect on their communities and lead to loss of confidence."

In recent years, the Research Councils have sought to make their processes more transparent. Their websites contain statements on their strategies for peer review, peer review procedures, the structure and membership of peer review panels and committees, and the membership of the Peer Review Colleges. The Working Group endorses the steps the Research Councils have tried to take in this direction.

We note that the Research Councils rely on peer reviewers to provide constructive comments, explaining their grades and views, which they can then send on to applicants. It is important therefore that peer reviewers take the time to provide helpful feedback. In addition, it can often be difficult to give constructive feedback to highly rated proposals that were unsuccessful, because of a shortage of funds. But in these circumstances Research Councils should not feel that they have

31 The report was published in September 2006 and is available from http://www.history.ac.uk/digit/peer/Peer_Review_report2006.pdf

to invent reasons for rejection other than 'we had to make a choice' and funding was limited. It may be that the Councils need to look at ways in which they can draw these issues more prominently to the attention of prospective applicants, in order to raise awareness of these difficulties and better manage expectations.

3.7 Conclusions

- (1) We have been impressed by the widespread support that the peer review system continues to receive in respect of grant awards. As with publications, the Academy sees no alternative to reliance upon peer review as a process integral to the maintenance of quality.
- (2) The Research Councils' efforts to make their procedures more transparent are welcome.
- (3) The system relies upon the willingness of individual academics to play their part. (see Chapter 5).
- (4) There is an inevitable element of independent judgement in peer review. This is not to be regretted but welcomed. In particular, we were impressed by the willingness of the AHRC Research Committee to exercise its judgement in respect of awards, and the willingness to act in this way is a strength of the AHRC system.

Chapter 4: Training and Prior Preparation for Peer Review Activities

4.1 Introduction

As we conducted our review, we were struck by the mismatch between the importance of peer review in the rhetoric of research policy and the lack of attention to understanding of it as a practice. In particular we were struck by the extent to which there is little attention to training in peer review.

Training is important, not just in itself, but because of the privileged position that peer reviewers enjoy. By virtue of reading a paper, reviewers can acquire access to original data sets, new empirical results or innovative conceptual work. In the business world, these would count as commercial secrets. In the academic world, the ethos is that reviewers are part of the gate keeping system the ultimate rationale of which is the fast and efficient dissemination of research findings. The integrity of the peer review system is therefore of great importance. One of the ways in which that integrity is maintained is through its dependence upon professional and unselfish motivations, and this in turn suggests the importance of training in the professional and ethical conventions of the practice.

4.2 Training Available to Postgraduates

So far as we can tell, there is little by way of formal training in peer review provided for those in the early or developing stages of their career. ESRC and AHRC training guidelines do not identify it as an essential element in the training of postgraduate researchers. In our discussions with postdoctoral fellows, there was a clear expectation that participation in peer review was an integral part of academic life, together with an acknowledgement that their own training had not prepared them for the tasks that were involved. They said that they had picked up the principles of the practice through informal apprenticeship from mentors and supervisors, together with reflection on peer review reports they had seen.

Given the general importance attached to peer review, it is important that new generations of scholars are inducted into its norms and values. We recommend that formal training in peer review and its principles be incorporated into the training guidelines of the Research Councils and HE institutions. This training should be orientated not only to inculcating a concern for academic quality, but also a recognition of the issues of professional ethics that are raised by peer review, including the importance of intellectual property rights and the need to deal fairly with the work of others.

In order to learn about peer review, graduate students of postdoctoral researchers need to have experience of peer review at first hand. The members of the Working Group identified different ways to achieve this goal. For some members, the only way in which training can be adequate is for journal editors to be willing to assign graduate students as well as junior postdoctoral researchers as reviewers, with appropriate guidance and mentoring. Those who held this view saw the main safeguard as lying in editorial control and judgement. For other members of the Working Group, assigning graduate students to review a paper (as distinct from postdoctoral researchers) raised questions about the extent to which any such assignment is consistent with the fundamental principle of peer review, namely that quality is judged by those with demonstrated competence. Ways around this dilemma include using graduate students and junior postdoctoral researchers as third referees, editors maintaining their independence of judgement and transparency with authors about the status of reports from early career researchers.

Journal editors are able to give feedback to those who undertake reviews by circulating copies of their decision letters together with the full set of review reports. It can often be helpful for those involved in the review process to see the basis upon which decisions have been arrived at. Where journal editors have the resources to adopt this practice, we commend it as a way of contributing to understanding for those who undertake peer review.

4.3 Training in Grant Awards

Because there are fewer alternatives in the case of grant applications than there are for journal submissions, it is important to ensure that the review process is as sound as it can be. If journal reviewing requires training, the same point is even more applicable to those engaged as peer reviewers for grant applications.

We welcome the fact that the development of AHRC and ESRC colleges of peer reviewers has allowed training to take place for peer reviewers. This is valued by those who work in the system. The AHRC panellists, in particular, spoke highly of the establishment in 2004 of the AHRC Peer Review College, and the induction days provided by the AHRC which enable panellists and reviewers to train alongside each other. As one panellist noted,

“This has not only encouraged a strong sense of collegiality, but also means that reviewers now have a better understanding of what is expected of them. The quality of the reports we get has improved dramatically.” (comment from AHRC panellist)

The Working Group believes that the AHRC training has been a model of good practice. Other Councils might wish to look at AHRC practice to see if there are any lessons that they can draw from it. The Group also commends the ESRC’s practice of giving feedback to its peer reviewers

4.4 The Contribution of Non-Academic Experts to Peer Review

We have already noted that the definition of peer review is not straightforward, in the sense that in many disciplines - particularly applied aspects of the social sciences - practitioners, public servants and social entrepreneurs may be important voices in determining whether a grant should be made or a paper published. Members of the Working Party who have sat on assessment panels involving practitioners know that practitioners can often be as insightful on purely academic issues as on implications and applications.

However, it is also important for practitioner and potential user reviewers to understand that intellectually innovative research is not like buying from a shopping catalogue: one cannot just specify the design and order the product to time. The image of consultancy is also remote from genuine research. Here again there is a role for training, so that those whose working life is not one of research can understand the logic of enquiry that underlies research and the need to protect its autonomous development.

Where non-academic reviewers are involved in peer review in the assessment of grant awards, we recommend that there be some training or communication to ensure that the academic integrity of the work is not compromised in a search for supposed relevance.

4.5 Conclusions and Recommendations

Because peer review relies upon professional norms, it is important that these are reinforced by appropriate training.

R1 Training guidelines should be amended so that explicit attention is given to the need to train postgraduates in the norms and conventions of peer review. Training should also be made available to early career track postdoctoral researchers, which could be an effective use of some of the ‘Roberts’ monies’.

For Research Councils UK (RCUK), Arts and Humanities Research Council (AHRC) and Economic and Social Research Council (ESRC), other research funders, universities and their representative bodies

We were impressed by the AHRC induction days, which enable panellists and reviewers to train alongside each other. The AHRC panellists we interviewed said that reviewers had a much better understanding of what was expected of them, and that the quality of their reports had improved markedly. We encourage other research funders to look at AHRC practice to see if there are any lessons that they can draw from it.

R2 Recognise the importance of training to ensure the competence and propriety of referees. We commend the AHRC programme of induction and training for its panellists and members of its Peer Review College.

For humanities and social science (HSS) research funders

The Research Councils are under pressure to employ non-academic peer reviewers. In many circumstances this is to be welcomed, but the practice highlights the need for training and clear communication.

R3 Give further consideration to the way in which non-academic expert users contribute to research council peer review assessments, with particular emphasis on the training and support that they receive.

For AHRC, ESRC and other research funders

Chapter 5: The Costs of Peer Review

“I am conscious that I’m asking someone to find several hours of their time to do something for me basically for free.” (comments from humanities and social science journal editor)

“Editing a journal... is sheer opportunity cost. You get less than no recognition for it from either your own department or the discipline more generally.” (comment from humanities and social science journal editor)

5.1 Motivations and Incentives

Two features stand out about peer review. First, editors and research funders often say how impressed they are by the quality of the reports that they receive from reviewers and the extent to which reviewers take trouble over their task. Second, there is little recognition or reward for peer review activities. It is by and large an anonymous activity. Some journals publish a list of names of those who have reviewed for them, but there is little public recognition of the scale of effort that is involved. Similarly, research funders such as the AHRC and ESRC publish the names of those who are members of their Peer Review Colleges, but reasons of confidentiality preclude their saying which referees were assigned to assess a proposal or set of proposals. Given these two features of the process, why do academics act as reviewers in the conscientious way that they do?

As part of our own work we held group discussions with those who undertake peer review. The range of motivations that emerge from our discussions varied, but typically related to the professional value and satisfaction that refereeing provides. Some respondents see refereeing as one of the responsibilities of being an academic, and as a way by which one can, as an individual, make an important contribution to the development of the subject. Others say that it is a way of keeping up with new developments, and that the activity is not all cost, since one learns things through undertaking the work. Still others see it as a form of reciprocity: one referees for others in return for being refereed. Box 5.1 gives a range of these reported motivations.

Box 5.1 The Range of Motivations for Peer Review

“In my experience, the large majority of referees have been very willing to referee content and are pleased to be asked. Often because they know the journal and what it is trying to achieve, but sometimes because they do not know the journal and it is an opportunity to find out more.” (from our survey of humanities and social science journal editors)

“a remarkable insight into developments in the discipline” (from our survey of humanities and social science journal editors)

“a growing network of colleagues and friends” (an incentive cited frequently in our survey as a reason for editing a humanities and social science journal)

“the capacity to shape the field in certain directions” (another frequently cited reason for editing a journal)

These accounts from our own conversations echo those that have been found by others who have investigated journal reviewing. Mulligan, for example, cites the responses from a focus group of reviewers convened by Elsevier. Among their reported motivations were a sense of duty, a recognition of reciprocity, a general interest in the subject and a desire to keep up to date with the subject. For younger scientists, it was thought an honour, confirming standing in the scientific community.³² There is in these accounts a mixture of professional commitment, a sense that reviewing is a responsibility and a recognition that something of value for the reviewer might also emerge.

With grants, motivations can be similarly complex. One additional incentive for reviewers of research grant proposals may be that they gain insight into funders' peer review processes, which may give them (or their colleagues) an advantage when they prepare proposals to that research funder.

5.2 Journal Review and Full Economic Cost

Given that peer review is central to the quality of published research, the question arises as to whether it should be more explicitly recognised and rewarded. At a time when a regime of full economic costs is being introduced into the funding of research, we can ask whether the same logic could be applied to the peer review process. What might it mean to pay full economic costs for someone's peer review activities?

To gain an idea of the numbers involved, let us work through a possible costing. Suppose that an average article in the humanities takes between four and five hours to read, evaluate and write a report on. Full economic costing would use an hourly rate to gain an idea of the financial sum that notionally should be charged for this work. Suppose that we work to a figure of £100 per hour (a conservative figure), and we assume that on average an article takes 4.5 hours to evaluate. Then each referee can be costed at £450 per article, and if there are two referees per article, then the cost per article is £900. Even if we allow for modest costs to an academic press of publishing an article - say £1000 per article - the costs of peer review nearly double those production costs. We stress that these costings are modest. For some subjects academics can charge much higher rates for consultancy: leading Economists can easily charge £3,000 to £4,000 a day for their services; and some papers take much longer to review.

Some of the large commercial journal publishers could afford to pay significant sums to individuals for their refereeing, but many journals in the humanities and social sciences are small scale, with print runs in the hundreds rather than the thousands. Such journals could not exist if they had to pay their referees.

Similar points can be made about the difficulty of journals in the humanities and social sciences moving to an 'open access' model, in which contributors paid to have their papers evaluated for publication and referees were paid for their time. Under this model, journals would have to charge submissions at a rate to cover their costs. For the natural sciences and medicine, in which the costs of research are paid for out of a grant, it is possible to imagine something like a full cost charging regime. By contrast, the small scale, often unfunded, research in the humanities and social sciences could not survive under such a model.

Moreover, there are real questions as to whether peer reviewers (or the institutions to which they belong) ought to be able to cover the cost of their time, particularly in humanities and social science subjects. Some journals do pay their referees amounts that are small compared with the actual cost, but this practice may be counter-productive because of the "crowding out" phenomenon

32 Adrian Mulligan, 'Is Peer Review in Crisis?', *Perspectives in Publishing*, 2, pp. 1-6, at p. 3 (see page 47 for full web page URL)

documented by Bruno Frey.³³ His evidence shows that systems may work less effectively after being bolstered by money payments, because introducing a money incentive tends to displace any code of honour that operated previously.

The Working Group concluded that FEC cannot be applied generally in the humanities and social sciences to the activity of peer review. Even where it could be used, as with some journals for which publishers could pay, it might be counterproductive to do so, because of the “crowding out” effect. Professional motivation remains the best guarantee of quality assessment.

5.3 The Costs of Peer Review for Grants

This issue has been the subject of a major review by Research Councils UK (RCUK)³⁴, assessing the costs of the peer review mechanisms that support research funding decisions.

The findings of the large-scale survey RCUK commissioned to assess the amount of time researchers and reviewers across all academic disciplines spent on research grant proposals suggested that the costs to the research community in preparing grant proposals was up to £186 million (see paragraph 1.2). As demand far exceeds supply, RCUK considered ways in which these costs could be reduced.

On the assumption that the size of the research budget is unlikely to increase to a level that will match the number of high-quality applications, three of the four options considered by RCUK focused on reducing the number of applications. Possible methods included: consolidating grants, introducing institutional level quotas, or controlling re-submissions. A fourth possible method focused on reducing the time taken to prepare applications by increasing the use of outline bids.

During the enquiry concerns were expressed within the humanities and social science research community that a regime of consolidated larger grants might lead to ossification, and to a further concentration of research funding in the hands of a small number of institutions at the expense of smaller pockets of research excellence. In addition, the consolidation of grants could work against efforts to reduce costs, and, even more importantly, to ensure that the most innovative research is funded. The consolidation of grants would probably increase the amount of effort that universities would put into each application, because the awards would be even more valuable than they have been in the past.

“The cost of the process is determined to a very great extent by how much is at stake for the applicant not by the extent of the formal data requirement. Simply trying to ration or discourage applications means that each application is more critical. This means that any fall in the number of applications will be offset by an increase in the cost per application.”³⁵

RCUK announced in June 2007 that it would only be consolidating grants into larger and longer awards, where it was appropriate for the research area, because it recognised that consolidation would pose problems for certain disciplines, namely, arts, humanities, social sciences and mathematics. The Working Group welcomes this acknowledgement. In disciplines where costs are relatively low, the ability to award small grants remain important. The Working Group also commends the efforts made by research funders to ensure that their peer review processes are proportionate to the scale of award, with smaller grants attracting less scrutiny, although there can be considerable differences as to what counts as small or large.

33 Bruno Frey, *Not Just for the Money*, Edward Elgar, Cheltenham, 1977.

34 RCUK *Response to the Project Report and Consultation on the Efficiency and Effectiveness of Peer Review*, June 2007.

35 *A Dangerous Economy: the wider implications of the proposed reforms to the UK Research Councils' peer review systems*, Tom Sastry, Higher Education Policy Institute, December 2006.

5.4 Conclusions and Recommendations

Surveys and other evidence have shown that there are various reasons why academics participate in peer review. Not all motivations are altruistic, and there is no reason why they should be. However, a central element, without which the peer review system would not exist, is the professional commitment to contribute to the academic public good. There is usually no practical way in the humanities and social sciences, in which the full economic cost (FEC) of peer review activity for the purpose of journal publication can be recovered. This reality needs to be recognised within the funding of research and universities.

By a similar token, the application of FEC would make the funding of most 'open access' journals in the humanities and social sciences impossible.

The significant contribution made by peer review to the health, well-being and development of academic disciplines needs to be better understood and recognised. Peer review has undoubtedly helped to underpin the success of UK research in the humanities and social sciences. The importance of these activities should be better reflected in research assessment exercises. Those responsible for the management of universities and research institutes need to ensure that they have processes in place to encourage and reward peer review activity, recognising that the benefits to the institution are indirect rather than direct.

R4 Each university in receipt of public funds should accept an obligation to encourage its researchers to engage in these activities, recognising that peer review is an essential part of the fabric of academic life - the costs of which are met by the funds allocated by the funding councils to support research infrastructure.

For University Vice-Chancellors and Universities UK (UUK)

Peer review will always impose costs, not least upon grant applicants, and the financial burdens of peer review can be measured in different ways. Any method that simply looks at the time-costs of peer review relative to size of awards is bound to cause problems in the humanities and social sciences, where awards are typically smaller than in medicine and the natural sciences. Policy makers need to be aware of these differences, and we commend RCUK³⁶ for taking note of them. Research Councils need to ensure that their standards of peer review, whilst robust in respect of quality, are proportionate in terms of the burdens they impose.

R5 Develop a more sophisticated understanding of the costs of peer review, and maintain efforts to ensure that the peer review processes are proportionate to the scale of award.

For RCUK/Funding Councils/relevant Government Departments

³⁶ RCUK Response to the Project Report and Consultation on the Efficiency and Value-for-Money of Peer Review, June 2007

Chapter 6: Metrics and Peer Review

“There is no absolute divide between metrics and peer review.” (Rt Hon Bill Rammell, Minister for Higher Education, June 2006)

“Metrics will not remove distortions and game-playing, merely provide different ones.” (Sir Howard Newby, Vice-Chancellor, University of the West of England, June 2006)

6.1 Introduction

The *primary* role of peer review is to assess individual grant applications and work for publication. However, as noted in Chapter 1, peer review has been extended to encompass broader evaluations of research performance, of which the most important in the UK has been the Research Assessment Exercise (RAE). Hitherto, the RAE has been undertaken approximately every five years on behalf of the UK funding councils, with the aim of assessing the quality of research at UK higher education institutions. There are RAE panels (now sub-panels) of peer reviewers for each broad subject area. These panels assess the quality of research conducted by departments or units of assessment. Because each panel looks at publications and grant awards that have themselves been peer-reviewed, the RAE can be thought of as a form of *secondary* peer-review, but conducted at a higher level of aggregation.

In December 2006, the Government announced that a new framework for assessing research quality would replace the RAE after the 2008 exercise, with the aim of reducing the costs and distortions of the current exercise. The Government gave the Higher Education Funding Council for England (HEFCE) the task of developing this new framework. In the natural sciences, assessment after the 2008 RAE will be based on a basket of statistical indicators (metrics), including bibliometrics, research grant income, and postgraduate student data. HEFCE has confirmed that research quality in the arts, humanities, social sciences, and mathematics will be assessed by a light-touch process, based on peer review (in the secondary sense) but informed by metrics. This approach will be applied in 2013 for the allocation of funding from 2014 onwards.

Since bibliometrics and measures of grant income themselves embody processes of peer review, there is a complex set of interactions between the practice of peer-review and the use of metrics in research assessment. The move to a new system of assessment raises particular challenges in the humanities and social sciences, because of the special features of research in these disciplines, and also because the metrics that are currently available are much less well adapted to the purposes of secondary peer review than those for medicine and the natural sciences. The funding allocated on the basis of the research assessment exercises (QR funding) constitutes a much higher proportion of the total dual support funding for the humanities (at around 85%) than that for medicine and the natural sciences (at around 50%). Care therefore has to be taken that the new framework will not have adverse consequences on the way in which humanities and social science research is assessed for the purpose of funding decisions.

6.2 What is meant by Metrics and Metrics-Based Research Assessment?

In the context of research assessment, a metric is a summary measure of performance expressed in a form that enables the performance of different researchers to be evaluated and compared. A simple example of a metric for researchers would be to count the number of their publications in a given amount of time (over a five year period say). However, metrics can only be reliable in evaluating

research if they seek to assess quality as well as quantity. So the type of metric that ought to be of interest to policy makers is some *quality adjusted* measure of performance, such as the number of publications in journals of a specified quality in a given amount of time or the number of citations an article receives in a specific period. The main intellectual problem confronting the development of metrics is the construction of measures that reliably and validly capture quality.

Satisfactory metrics are of interest to policymakers for a number of reasons.

- (1) Metrics potentially allow for greater accountability for the monies that have been allocated for research, since they form the measure of output against which the input (the money) can be compared.
- (2) International benchmarking (the assessment of the quality of work of scholarly communities in different countries) is easier with metrics.
- (3) The costs and complexity of the RAE could be reduced, since panels conducting their secondary peer review would have summary measures available of the quality publications and research grants.

Can satisfactory metrics be developed however? In this chapter we look at two approaches.

6.3 The Citations Approach

Citations can be counted, and one way of measuring the quality of a publication to assess whether it appears in a high-citation outlet. For example, it is possible to ask whether an article appears in a journal with a high 'impact factor'. The chain of inference involved here runs as follows. The more papers in the journal are cited, the more impact that journal has. The more impact a journal has, the more authors will want to publish in that journal. The more authors who want to publish in the journal, the more demanding will be the selection criteria applied in the refereeing process. The more demanding the selection criteria applied in the refereeing process, the better the average paper will be. The better the average paper in the journal, the more it will be cited. And so a virtuous circle is completed.

It is possible to imagine that an RAE sub-panel could supplement, or even replace, its own secondary peer-review, based upon its reading and judgement of submitted work, by metrics related to citation measures, controlling for impact factor. In adopting such an approach, a panel would be drawing upon the judgements of members of the academic community at large as reflected in peer review judgements of decisions to publish, and the citation rate of papers published. One advantage this approach would have is that it would make the assessment of the quality of research for any researcher in a unit of assessment less dependent upon the judgement of a single panel member. However, there are considerable difficulties in the way of implementing this approach for the humanities and the social sciences.

- (1) Social science and humanities journals typically publish fewer articles than journals in medicine and the natural sciences. In consequence, journal rankings are sensitive to small number problems and can move erratically and significantly from year to year, often reflecting citations in relation to only one article. Whilst it can be argued that over time journal rankings will remain stable, there is also movement in those rankings (and it is highly desirable for incentive purposes that there should be) and inferences from ranked journal to author quality are not easy to make.

- (2) Assessment of impact is complicated by the problem of time scale. The Thomson Scientific measure of two years reflects practices in the natural sciences and medicine, where specific results are picked up quickly in the literature. However, in the humanities and social sciences important papers may require greater time to be understood and absorbed than is true of papers in medicine and the natural sciences. Particularly striking examples of this phenomenon are to be found in the biographies of Nobel Prize winners in Economics. It took thirty years for the work of John Nash in game theory to be appreciated. The same is true of William Vickrey on tailor-made auctions of public assets.
- (3) To be eligible for a citation ranking a journal needs to be included in the Thomson Scientific journals list. The Working Group acknowledges that Thomson Scientific has good reasons for being selective in its journal inclusion, since the existing criteria of selection involve considerations of quality. For example, timeliness of publication is important, since “it implies a healthy backlog of manuscripts essential for ongoing viability.”³⁷ Other criteria may raise more difficult questions however. For example, journals in the humanities and the social sciences may have a distinct cultural or area orientation, so that such a journal would find it difficult to meet tests about the international diversity of its authorship. The more extensive the use of citation measures, the more important become the selection criteria for journal inclusion; this area needs closer scrutiny.

The Working Group does not conclude from these difficulties that there can be no role for citation measures among the metrics of research assessment, but warns that there can be no easy assumption that forms of metrics suitable to the natural sciences and medicine can be transposed to the social sciences and humanities. It endorses the general stance already taken by the British Academy that if metrics could be made to work, they could form a useful supplement to the secondary peer review currently represented by the RAE.³⁸ It also echoes the view of the Expert Group, set up jointly by the AHRC and HEFCE to explore the use of metrics in the assessment of arts and humanities research, which stated that:

“There is no single metric that is appropriate to measuring the performance of arts and humanities research. Of the metrics available, some are well established, while others are being developed. The metrics chosen as part of the assessment framework should reflect the multiplicity of peer review systems which are already in place and are an integral feature of academic life.”³⁹

However, the Working Group recommends that before metrics are used, pilot and exploratory studies should be undertaken (for example of the robustness of evaluations to different ways of measuring impact) to test the approach, especially in small or otherwise vulnerable subjects.

6.4 Monographs and Metrics

For many humanities and social science disciplines, the monograph remains the principal and most highly valued method of scholarly publication. Even within the present system of RAE assessment, it is not clear whether the quality of monographs is being properly assessed. The present

37 James Testa, ‘The Thomson Scientific Journal Selection Process’ at <http://scientific.thomson.com/free/essays/selectioofmaterial/journalsselection/>, p. 2 (last accessed 21 March 2007).

38 *Response to the DfES Consultation on the reform of higher education research assessment and funding*, October 2006 available from www.britac.ac.uk/reports

39 *Use of research metrics in the arts and humanities*: Report of the Expert Group set up jointly by the Arts and Humanities Research Council and the Higher Education Funding Council for England, October 2006.

system weights journal articles and monographs by the same value, although there are in most cases very different degrees of effort involved in their production, with monographs taking longer to produce and often embodying many years of research.

The move towards metrics-based assessment makes more sense in subjects in which journals are the principal method of publication, because the citation indices calibrate articles by their journal citations and journals by their impact factors. Although citation indices do capture citations to monographs, they do not capture citations within monographs. We understand that it is technically possible to do this with electronic publishing. As metrics are developed, the Working Group recommends that as a priority work should be commissioned to assess whether the absence of monograph to monograph citation assessments is likely to pose serious problems for the humanities and social sciences.

6.5 The European Reference Index for the Humanities (ERIH)

The problems of citation analysis discussed above has led the European Science Foundation (ESF) to explore an alternative model in its project to produce a European Reference Index for the Humanities (ERIH). In particular, in response to concerns that existing citation indices too rarely included the best journals published in languages other than English, the ESF's Standing Committee for the Humanities (SCH) has sought to compile its own database of scholarly journals in the humanities.

The ERIH has four self-proclaimed aims:

- (i) to increase visibility for humanities research in Europe
- (ii) to disseminate European research in the humanities worldwide
- (iii) to encourage best practice in the publication of humanities journals
- (iv) to provide a benchmarking tool for comparisons at aggregate level

In pursuit of these goals, the ERIH has produced listings of journals, broken down into three categories: A, B and C. The definitions of these categories are as follows:

Category A: high ranking international publications with a very strong reputation among researchers of the field in different countries, regularly cited all over the world.

Category B: standard international publications with a good reputation among researchers of the field in different countries.

Category C: research journals with an important local/regional significance in Europe, occasionally cited outside the publishing country though their main target group is the domestic academic community.

Details of the project can be found at www.esf.org/research-areas/humanities/activities/research-infrastructures.html

It has been stressed that in producing its classification of journals, the ESF is not seeking to make or presume judgements of quality but rather to identify the scope and general interest of different journals, distinguishing between those that were highly specialist and those that were general.

“The distinction between categories A, B and C is not primarily qualitative; rather, the categorisation is determined by a combination of characteristics related to scope and audience. Journals have different profiles, and their scope of audience and field vary. Papers in journals with wide international prestige are not automatically of higher quality than papers in journals which are known and read only in a very specialised field.”(ESF/ERIH website)

We understand the desire of those producing the ERIH to overcome the limitations of existing citation approaches, but as we have looked at the ERIH’s own approach, we have identified a number of problems. Some of these problems are matters of implementation, which could be overcome with further development, but other problems raise questions about the conceptual foundations of the approach:

- (1) The statement on the ESF website (quoted above) to the effect that the classification is not intended as a ranking of quality is incompatible with the way in which the classes are defined. For example, journals in the A category are said to have ‘a very strong reputation’ whereas those in the B category have ‘a good reputation’. It is difficult to read these definitions without taking them as quality rankings.
- (2) It is not clear what the relationship is between citation information and ranking by putative journal esteem. The criteria refer to journals on the A list as being ‘regularly cited all over the world’, but there is no independent citation information given to support this assessment. It would have been helpful if the classifications had been systematically related to existing citation measures, so that a comparison could have been made between the two approaches.
- (3) Although the panels devising the classifications were made up of scholars distinguished in their own field, the number of such scholars is small. There are techniques of journal classification that use the judgements of a much larger number of people. For example, Garand and Giles have set esteem measures alongside citation measures for journals in Political Science.⁴⁰ The advantage of using the judgements of a relatively large number of people is that individual errors in classification are cancelled out.

We were concerned lest our own response to the ERIH initiative reflected a narrow set of methodological pre-occupations, so we asked Academy Fellows for their views of the ‘initial lists’ for their disciplines. Box 6.1 gives some examples of the comments received. These comments suggest that the methodological worries that we have complement the substantive concerns of those who are expert in specific disciplines and sub-disciplines.

⁴⁰ James C. Garand and Michael W. Giles, ‘Journals in the Discipline: A Report on a New Survey of American Political Scientists’, *PS: Political Science and Politics* 36: 2 (2003), pp. 293-308.

Box 6.1: Comments from Fellows regarding the ERIH 'initial lists' for their disciplines

"Especially problematic in this connection is the fact that Category C in this exercise is different in kind from A and B in that it is devoted to 'journals of local and regional significance in Europe'. If it is felt necessary to distinguish a category of journals in this way, it would be better to give them a different type of label, perhaps calling them simply 'Local' rather than C. In this way their distinct status is recognised but there is no implication of hierarchy with respect to journals in the other two categories.

This is important because often the journals in the C list are the leaders in their field. To take an example, *Studi di Filologia Italiana* has been in existence for more than half a century and has numbered some of the most eminent scholars in Italian textual philology and historical linguistics amongst its editors and contributors, including many from outside Italy. Inevitably, however, Italians figure most prominently since it is - not surprisingly - within Italian universities that this discipline finds a particular concentration of research excellence. And for this it is rewarded with a classification as C!" (comments from Fellows of the Academy's Linguistics Section.)

"I have major problems with the criteria: they relate more to language and distribution than to the intrinsic quality of the contents of these journals. So if the journal is in English, or predominantly so, and in consequence has the possibility of wide circulation and wide citation it has a higher chance of getting A than if it is not in English. But this has really nothing to do with quality. A case in point is *Tarbiz*. This is in Hebrew so its readership is limited, but among core researchers in the field of Jewish Studies it is a major journal and should be in the highest category." (comments from Fellows of the Academy's Theology and Religious Studies Section.)

"There is a clear preference to rank English journals as A but no measures of impact or scholarly standing are offered in support. Moreover, nearly all the US Journals listed here receive an A grade. This seems to go against the stated aim which was to question the ISI bias towards US/English journals." (comments from Fellows of the Academy's Archaeology Section.)

"An anomalous consequence of the ERIH methodology is that the same journal may get two or more different rankings within different subject lists, if it is deemed to be 'central' in one field and 'peripheral' in another..." (comments from Fellows of the Academy's Linguistics Section.)

"Continental musicology is a much more narrowly defined and traditional field than its British and North American equivalent.... as a result, the categorisations in this list [Musciology] reflect what most British and American musicologists would consider an entirely outmoded concept of the discipline." (comments from Fellows of the Academy's History of Art and Music Section.)

We conclude that the European Reference Index for the Humanities as presently conceived does not represent a reliable way in which metrics of peer-reviewed publications can be constructed. We suggest that alternative approaches are needed to deal with the problems. These alternative approaches might include the following.

- (1) Analysis could be conducted of journals that are not currently included in Thomson Scientific's citation indices, but are nonetheless cited in the journals that are included. These cited-only journals may provide evidence of quality that is neglected in current measures.
- (2) Academic communities could be consulted widely about their perceptions of journal quality using electronic communication. There is no reason why such an approach should be regarded as an alternative to the classification produced by a small group of experts, but it would provide some comparative reference point of assessment.

6.6 Conclusions and Recommendations

Metrics and peer review are related in complex ways. This partly reflects the extent to which peer review describes many diverse practices. Metrics are typically summary measures of peer reviewed activities, and the use of metrics sets up incentives to change behaviour in ways that affect peer review.

R6 Care should be taken to ensure that any metrics employed reflect the distinctive nature of the humanities and social sciences research and do not have an adverse affect on the quality of the work that they are seeking to measure. HEFCE should commission an investigation to explore whether there is scope to modify the Thomson Scientific indices to accommodate the special features of humanities and social science research. Until this work is completed and assessed, metrics should remain an adjunct to the research assessment (RAE) panel peer review process rather than a substitute.

For Higher Education Funding Council for England (HEFCE)/RCUK/relevant
Government Departments

Difficulties with existing metrics have led to attempts to develop alternative approaches, notably the European Reference Index for the Humanities (ERIH) that involves the ranking of journals. However, the Working Group identified a number of methodological difficulties with the ERIH approach, relating to reliability and validity, some of which were fundamental.

R7 The experience of the European Reference Index for the Humanities (ERIH) shows how difficult the task is, and we conclude that the ERIH does not at present represent a reliable way in which summary measures of peer reviewed publications can be constructed.

For Higher Education Funding Council for England (HEFCE)/RCUK/relevant
Government Departments

Chapter 7: Peer Review and Innovation

7.1 Introduction

We have already noted that a common point of complaint in connection with peer review is that it has a bias against innovation. Authors sometimes say that peer review has blocked publication of their most innovative work. Grant applicants claim too that their innovative proposals have been blocked, and at various times the Research Councils themselves have expressed concern on that score.

Both complaints are serious in themselves, but they have recently been augmented by new concerns that peer review is a barrier to knowledge transfer and therefore innovation in the economy and society at large. From this point of view, it is argued, peer review needs to be adapted in order to meet the challenges that the public interest poses.

Clearly these are different levels of criticism and the possible remedies are correspondingly distinct. The concern with innovation and knowledge transfer is something that may be aimed at a task that peer review in its conventional form is not designed to deal with. However, if peer review prevents the publication of intellectually innovative work or blocks innovative research proposals, then the potential indictment is more fundamental.

7.2 Innovation and Publication

There is no lack of evidence that peer review can undervalue innovative intellectual work and lead to the delay or displacement of important papers. For example, Gans and Shepherd relate the experiences of a large number of Economists who have either won the Nobel Prize or the John Bates Clark Medal, who nonetheless have had their papers rejected by peer review journals⁴¹ These papers include items that, when they were eventually published by a second or subsequent choice journal (or even in a collection of essays), turned out to be classics. They include Akerlof's 'The Market for Lemons' as well as papers by Gary Becker, Jagdish Bhagwati, Roy Harrod, Robert Lucas, Tibor Scitovsky and James Tobin. Other experiences are also telling. Karl Popper famously had his initial draft of *The Open Societies and its Enemies* repeatedly rejected. John Nash's suggestion that Nash equilibrium admits an evolutionary interpretation was equally famously deleted from his Nobel Prize winning paper on the grounds that it was of "no interest".

We can only speculate on the broader causes of why innovative papers are rejected, but they presumably include the tendency of editors to play safe and weigh negative judgements more heavily than positive ones. No doubt too the existence of informal communities of scholars working in specialist sub-fields and friendly towards work of their own kind and hostile to other sorts of work also plays a part. Finally, perhaps the shock of the new itself is a barrier to publication. In the Gans and Shepherd sample, many of the respondents say it is their most original papers that are rejected.

However, although there are these extremely telling examples, the evidence they provide needs to be counterbalanced by other considerations.

- (1) In a number of cases, the authors of the rejected work sometimes recognised that where they had received rejections, the decision was justified. Despite their general eminence and achievement, even the members of this group were capable of producing work of variable

⁴¹ Joshua S. Gans and George B. Shepherd, 'How Are the Mighty Fallen: Rejected Classic Articles by Leading Economists', *Journal of Economic Perspectives*, 8: 1, pp. 165-79.

quality, and the barrier of peer review saved the embarrassment of publishing a piece of which, in retrospect, they would not have been proud. Not everything that is written by an innovative researcher is itself innovative research.

- (2) Although the study revealed that in some cases the researchers were discouraged from publishing a paper at all, there was also a quality-enhancing effect from work that did get through the peer review process, and this was acknowledged by some respondents in the study.
- (3) Rejection by one journal does not preclude acceptance by another. One of the strengths of the peer review process is this decentralised quality, which means that decision-making is dispersed and non-cumulative. In our view this is another reason to avoid a rigid classification of journals of the sort proposed by the ERIH.
- (4) Any quality control system will produce false negatives, but the existence of false negatives is a reason for maintaining editorial independence. In a situation in which editors simply go with the opinion of referees, they may miss the opportunity to advance the subject by accepting for publication the unusual and the innovative.
- (5) It is important not to commit the fallacy of assuming that, because high quality will be innovative, the innovative is necessarily high quality. There is a need to affirm that there are other criteria of assessment, and sometimes these will be in conflict with a concern for novelty. These other criteria include: accuracy, validity, replicability, reliability, substantively significant, authoritative and so on.

7.3 Innovation and Grants

The Boden report drew attention to the intrinsic difficulties of keeping abreast of new research areas, capturing new research talent, and spotting novel and innovative work. While it is difficult to find evidence that innovative projects receive less than fair treatment, low success rates, together with the move towards larger and longer grants, may make it more difficult for some Research Councils to make the kinds of risky, speculative decisions that are often necessary to support the most innovative research. We therefore commend the AHRC's grant scheme for speculative research, which aims to "support high-quality research where the speculative, experimental or exploratory nature of the work means that results or outcomes are uncertain or cannot be guaranteed, or where a significant degree of risk is involved."⁴²

We also commend a balance between speculative, high-risk research and safer work that develops ideas already well established in response mode funding. The use of small pump priming grants can be very effective in this context, especially those that are targeted at early career track researchers and those with novel approaches. The British Academy itself makes a large number of small grants each year aimed at providing pump priming for projects, many of which subsequently go on to be successful and obtain higher levels of funding from bodies like the ESRC and AHRC. Similarly, a number of universities offer small pump priming grants to their staff with these aims in mind, and often support these grants by drawing on their endowment funds. This seems to be a sensible way to use these funds, and there may be scope for the Government to consider ways in which they can encourage endowments of this kind within universities to support small grants for innovative, high risk research.

⁴² The Route for Speculative Research - www.ahrc.ac.uk (see page 47 for full web page URL).

With regard to high-risk research, it is also important to understand that the success rates - in terms of outcome and impact of the research funded - may be low: nine out of ten projects of this kind may fail, but the investment will be fully justified if the tenth project proves to be a major success. It is important, therefore, to recognise and account for these risks.

Sometimes the Research Councils will identify strategic research initiatives (issues of national importance, requiring targeted funding). While funders may take soundings from both the research community as well as the 'user' community, we consider that the identification of these topics should also needs scrutiny. One way forward would be for such initiatives also to be subject to robust peer review procedures.

7.4 Knowledge Transfer and Impacts

Sometimes when referring to innovation, critics of peer review have something broader in mind than barriers to innovative academic work. They are referring to knowledge transfer and the idea that economic growth - and perhaps social welfare more generally - can be improved by the application of up-to-date research. In a knowledge-based economy, such concerns are increasingly common.

As a consequence, the Research Councils are coming under increasing pressure to demonstrate the economic benefits of the research they support. The Warry report, published in 2006, recommended that applicants for Research Council responsive mode grants "should identify potential economic benefits (if any) and reviewers should have clear guidance on how to score these benefits."⁴³ When submitting applications, researchers will be asked what benefits might arise from the research. RCUK has made it clear that it will use a broad interpretation of benefits, which will include the contributions to social and cultural well-being. This has gone some way to allaying concerns that judgements of quality may be distorted by consideration of potential impact. However, other concerns remain.

- (1) It is inherently difficult to predict at the outset of a project what likely benefits the research might bring. Even when research has been completed and published, it may be many years before its impact can be properly understood and assessed. As an earlier Academy review, *"That full complement of riches": the contributions of the arts, humanities and social sciences to the nation's wealth*, said:

"and how can we decide what is useful knowledge and what is not? We fear that those who try to second guess what is 'useful', and what is less useful, knowledge may make the wrong decisions. Predicting which areas will prove to be the most important and valued in the future is in its nature difficult, perhaps impossible. ...The outcomes of original research can never be predicted in advance, and if they could there would be little point in research in the first place."
- (2) The proposition does not seem to recognise fully the relationship between basic and applied research. Research seen initially as blue skies might have business applications. Applied research builds on the foundations laid by basic research. Both have to be funded properly, if the UK research base is going to prosper.
- (3) Too great a focus on knowledge transfer might distort research priorities, and the direct/indirect benefits flowing from research universities.

⁴³ Research Council Economic Impact Group, *Increasing the economic impact of Research Councils; advice to the Director General of Science and Innovation*, DTI, July 2006.

It is therefore a mistake to confuse the prospective benefits of humanities and social science research with the idea that there is a simple set of policy techniques to transfer good work in these disciplines into ways of enhancing social welfare. In part this is because the production of knowledge itself depends upon allowing the autonomous development of disciplines untrammelled by concerns for where that development is leading. In part it is also due to the impossibility of predicting the innovative. There is a need for the 'virtue of indirectness', so that we avoid the fallacy of thinking that there is a simple linear relationship between good research in the humanities and social sciences and innovative applications in ways that can be administratively managed. Below are two examples of research projects that had unforeseen practical implications that were not predicted in advance of their awards being made.

The ESRC Centre for Economic Learning and Social Evolution (ELSE) at University College London was established to undertake basic research on game theory. Research supported by the ESRC enabled researchers to develop auction theories, which were implemented at a government auction of mobile phone licences. It has been estimated that the figures generated were thirty times higher than the government had originally anticipated, netting the Treasury £22.4 billion in April 2000.

In 1997, a project was initiated at the University of Oxford by the Centre for the Study of Ancient Documents (Professor Alan Bowman) in conjunction with the Department of Engineering Science (Professor Mike Brady), to develop a new computer-based image-enhancement technique to decipher ancient stone inscriptions and writing tablets. The technique they developed broke new ground in the field, but also had other unforeseen benefits as members of the team of engineers realised that the technique could be used to improve methods of analyzing mammogram images for early detection of breast cancer.

7.5 Conclusions and Recommendations

Allegations are often made that peer review inhibits innovation. Some evidence exists that this can happen, but the problem needs to be kept in proportion. Any form of quality control will produce false negatives. We have found no compelling evidence that peer review systematically operates to inhibit innovative research.

In the broader sense in which innovation involves knowledge transfer, we recognise that the involvement of non-academic expert reviewers as peers can be valuable and sometimes essential. However, this should not lead to a dilution of the academic quality of the work. Such peer reviewers may also need training. One important safeguard for innovation is the decentralised nature of the peer review process. It is important to retain this, and it is equally important that the development of metrics does not create incentives towards a hierarchy of journals, such that the decentralised nature of the system is corroded.

The variety of practices in the conduct of peer review among journal editors in the humanities and social sciences is a strength, not a weakness. There is no gold standard of peer review practice that all should follow, and editors should choose a system of peer review that best suits their journal. Nevertheless, there are principles that should be respected, and wider dissemination of the variety of

practices and an understanding of their strengths and weaknesses would help editors, editorial boards and the wider community, both academic and non-academic.

R8 Uphold the editorial independence of journal editors.

For publishers and learned associations

Research funders should take pains to avoid a mechanistic approach in their decision-making processes for the award of research grants in the humanities and social sciences. One example of good practice in this regard are the AHRC subject-specific panels. As with journal editors, AHRC panels are able to reach independent assessments of quality and can treat reviewers as advisers rather than judges. Each proposal is considered on a case by case basis, avoiding a formulaic averaging of grades, so one low mark will not automatically rule out a proposal. Similarly, the ESRC processes for its standard grants are designed with safeguards to handle proposals receiving divergent reviews and marks. These approaches also help to ensure that intellectually innovative proposals, where there is likely to be a marked contrast in the views expressed by peer reviewers, can be assessed on their merits.

R9 Consider grant proposals on a case by case basis, taking pains to ensure that award decisions are not made by the application of formulae.

For HSS research funders

Issues of knowledge transfer and impact play an important role in public policy, and are likely to become more, not less, important over time. These issues surface in different ways: in relation to those who are designated peers (see paragraph 5.4 re use of non-academic users in peer review), to the criteria by which peers are asked to judge a proposal or piece of writing and to the mechanisms that are used to undertake peer review. Quality should not be sacrificed in favour of relevance and impact. Applied research ought to meet the same standards of research design, sample selection and evidential inference that applies to any sort of work (allowing for the practical difficulties of conducting applied research). Indeed, if research is being used by policy makers to take decisions on matters that have a direct effect on the quality of citizens' lives, the standards ought to be as high as possible.

Similarly, novelty cannot be regarded as a substitute for quality. Concerns are regularly raised about the extent to which peer review introduces a conservative bias into decisions on funding or publication. But it is difficult to find firm evidence to support it, particularly in relation to the humanities and social sciences. However, to the extent to which it is a problem, there are ways in which a counter-bias can be introduced.

R10 Ensure that considerations of applicability and relevance do not compromise judgements of quality.

For HSS research funders

We commend the AHRC for seeking to ensure that innovative proposals are dealt with through funds that explicitly recognise a degree of risk. Other research funders for the humanities and social sciences may wish to consider adopting the same practice.

R11 Set aside funds for risky, speculative projects.

We also commend a balance between speculative, high-risk research and safer work that develops ideas already well established in response mode funding.

For HSS research funders

R12 Ensure that there is a healthy balance between strategic funding and responsive mode projects.

For HSS research funders

The use of small pump priming grants for innovative research in the humanities and social sciences can be very effective, especially those that are targeted at early career track researchers and those with novel approaches. Many of these grant holders may subsequently go on to obtain higher levels of funding from other sources. A number of universities offer small pump priming grants to their staff with these aims in mind, and often support these grants by drawing on their endowment funds. This seems to be a sensible way to use these funds, and there may be scope for the Government to consider ways in which it can encourage endowments of this kind within universities.

R13 Encourage endowments within universities to support small grants for innovative, high risk research.

For UK Government, UK universities, UUK and other representative bodies for universities

The selection of topics for strategic initiatives is not straightforward. The identification of the main challenges facing the UK and beyond, as well as the ways in which the subject disciplines promoted by funders can best respond to the issues they raise, should themselves be subject to robust peer review procedures.

R14 Ensure that the process of selecting topics for strategic initiatives is also subject to peer review.

For HSS research funders

APPENDIX 1

How the Review was Conducted

Working Group Membership

The British Academy appointed a Working Group to oversee the work of the Review. Under the chairmanship of Professor Albert Weale, FBA, the Group met on a regular basis between June 2006 and July 2007 to oversee the direction of the review. Its members were appointed by the British Academy and are drawn from a range of subjects in the humanities and social sciences.

Chairman

Professor Albert Weale, FBA	Professor of Government	University of Essex
-----------------------------	-------------------------	---------------------

Humanities

Professor Marianne Elliot, FBA	Director, Institute of Irish Studies and Professor of Modern History	University of Liverpool
--------------------------------	--	-------------------------

Professor Marian Hobson, FBA	Professor of French	Queen Mary, University of London
------------------------------	---------------------	----------------------------------

Professor Nicholas Jardine, FBA	Professor of History and Philosophy of Science	University of Cambridge
---------------------------------	--	-------------------------

Social Sciences

Professor Robert Bennett, FBA	Professor of Geography and Chairman of the Research Committee	University of Cambridge
-------------------------------	---	-------------------------

Professor Kenneth Binmore, FBA	Professor of Economics	University College London
--------------------------------	------------------------	---------------------------

Professor Howard Glennerster, FBA	Professor Emeritus of Social Administration	London School of Economics
-----------------------------------	---	----------------------------

Secretariat

Ms Vivienne Hurley	Secretary	Working Group
--------------------	-----------	---------------

Sources of Information

The British Academy addressed the questions posed by the study by:

- (1) consulting key national organisations. We met senior staff from: Arts and Humanities Research Council (AHRC); Economic and Social Research Council (ESRC); The Leverhulme Trust; the Joseph Rowntree Foundation; Association for Learned and Professional Society Publishers (ALPSP); Institute of Historical Research.
- (2) canvassing the views of AHRC and ESRC panellists and peer reviewers.
- (3) gathering information from 'fact-finding' meetings with journal editors and postdoctoral researchers.
- (4) surveying journal editors in the humanities and social sciences.
- (5) seeking the views of the Academy Fellowship.
- (6) reviewing existing literature.

Survey of Journal Editors

The survey sought to identify current practices of peer review for journals in the humanities and social sciences. It was an on-line survey available through the Academy's website. A total of 96 journal editors completed the questionnaire. As the questionnaire was distributed partly by the Academy and partly by journal publishers, it is difficult to determine the response rate with any precision.

The majority of the respondents were from journals primarily based in the humanities, although there were several that were multidisciplinary, straddling the boundaries for the humanities and social sciences. A significant proportion of the journals (at 38 per cent) were learned society ones. The majority of the journals (87 per cent) were directed at an academic audience. The number of papers they received varied considerably, ranging from 8 to 580 per annum. More information about the survey findings is available on the Academy's website - www.britac.ac.uk/reports

Acknowledgements

The Working Group is grateful to all those who contributed to its enquiry. It would also like to thank the following individuals, who made some constructive comments on the penultimate version of this report. They are:

Ian Broadbridge	AHRC
Professor Robert E. Goodin, FBA	Australian National University
Gary Grubb	ESRC
Gail Lambourne	AHRC
Professor Tony McEnery	AHRC
Victoria Waite	HEFCE
Professor Michael Worton	Vice Provost, University College London

Appendix 2 - List of Background Documents and Sources of Information

Advisory Board for the Research Councils (ABRC), *Peer Review: A report to the Advisory Board for the Research Councils from the Working Group on Peer Review*, 1990.

Arts and Humanities Research Council (2006). 'Peer Review and the Evaluation of Digital Resources for the Arts and Humanities'. (online) available from:

http://www.history.ac.uk/digit/peer/Peer_Review_report2006.pdf (accessed 3 January 2007)

Arts and Humanities Research Council and the Higher Education Funding Council for England, *Use of research metrics in the arts and humanities*, 2006.

Arts and Humanities Research Council, 'The Route for Speculative Research' (online) available http://www.ahrc.ac.uk/apply/research/research_grants/the_route_for_speculative_research.asp (accessed 13 February 2007)

Blackwell, *Best Practice Guidelines on Publication Ethics: a Publisher's Perspective*, Blackwell Publishing Ltd, 2006.

Donohue, John J. III and Levitt, Steven D., 'Measurement Error, Legalized Abortion, the Decline in Crime: A Response to Foote and Goetz (2005)', Federal Reserve Bank of Boston, *Working Papers*, No. 05-15, January 2006.

Donohue, John J. III and Levitt, Steven D., 'The Impact of Legalized Abortion on Crime', *The Quarterly Journal of Economics*, CXVI: 2 (2001), pp. 379-420 at p. 414.

Fifield, Anna and Cookson, Clive, 'Seoul Searching: Koreans find their Rapid Development has Hard Scientific Limits', *Financial Times*, Thursday January 19 2006, p. 15.

Foote, Christopher L. and Goetz, Christopher F., 'Testing Economic Hypotheses with State-Level Data: A Comment on Donohue and Levitt (2001)', Federal Reserve Bank of Boston, *Working Papers* No. 05-15, available at <http://www.bos.frb.org/economic/wp/index.htm>.

Foerster, Thomas von, 'The Future of Peer Review', American Academy of Arts and Sciences, 2001.

Frey, Bruno, *Not Just for the Money*, Edward Elgar Publishing Ltd, Cheltenham, 1977.

Gans, Joshua S. and Shepherd, George B., 'How are the Mighty Fallen: Rejected Classic Articles by Leading Economists', *Journal of Economic Perspectives*, 8.1, pp.165-79

Garand, James C. and Giles, Michael W., 'Journals in the Discipline: A Report on a New Survey of American Political Scientists', *PS: Political Science and Politics* 36:2 (2003), pp. 293-308

Hames, Irene, *Peer Review and Manuscript Management in Scientific Journals: Guidelines for Good Practice*, Blackwell Publishing Ltd, 2007

King, Gary, 'Replication, Replication', *PS: Political Science and Politics*, 28:3 (1995) pp.444-52.

Lock, Stephen, 'A Difficult Balance. Editorial peer review in medicine', *ISI Press*, Philadelphia, 1985.

Merton, Robert and Zuckerman, Harriet, 'Patterns of Evaluation in Science: Institutionalisation, Structure and Functions of the Referee System', *Minerva*, 9: 1, 1971.

Mulligan, Adrian (2004). 'Is Peer Review in Crisis?', *Perspectives in Publishing*, 2, (online) available from: www.elsevier.com/framework_editors/pdfs/PerspPubl2.pdf. (accessed 17 April 2007)

Research Council Economic Impact Group. *Increasing the economic impact of Research Councils: advice to the Director General of Science and Innovation*, DTI. July 2006.

Research Councils UK. *RCUK Response to the Project Report and Consultation on the Efficiency and Effectiveness of Peer Review*, June 2007.

Research Information Network (RIN), Department of Trade and Industry (DTI), and Research Councils UK (RCUK). *Analysis of data on scholarly journals publishing*. 2006

Rowland, Fytton, Report Commissioned by the JISC Scholarly Communications Group, *The Peer Review Process*

Royal Society, *Peer Review - An assessment of recent developments*, 1995.

Sastry, Tom, *A Dangerous Economy: the wider implications of the proposed reforms to the UK Research Councils' peer review systems*, Higher Education Policy Institute, 2005.

Sense about Science, *I don't know what to believe: making sense about science stories*, 2005

Sieber, Joan 'How can we research peer review', *Nature*, 2005.

Smith, Richard, 'Opening Up BMJ Peer Review', *British Medical Journal* 318 (7175), pp. 4-5.

Testa, James, 'The Thomson Scientific Journal Selection Process' at <http://scientific.thomson.com/free/essays/selectioofmaterial/journalselection/>

The Association of Learned and Professional Society Publishers (ALPSP), *Scholarly Publishing Practice: Academic journal publishers' policies and practices in online publishing*, 2005.

The British Academy
10 Carlton House Terrace
London
SW1Y 5AH

Telephone: 020 7969 5200
Fax: 020 7969 5300
Email: chiefexec@britac.ac.uk

Web site: www.britac.ac.uk

© The British Academy 2007