

TOPIC:

Peer Review:

**Making Decisions
about Science and Technology -
between the Devil and the Deep Blue Sea?**

John Rigby

Some Developments in Peer Review

Michael Stampfer

Nr. 21

June 2004

BUNDESMINISTERIUM FÜR BILDUNG, WISSENSCHAFT UND KULTUR,
Minoritenplatz 5, A -1014 Wien
Mag. Markus Pasterk markus.pasterk@bmbwk.gv.at
Ursula Suda ursula.suda@bmbwk.gv.at

bm:bwk

BUNDESMINISTERIUM FÜR VERKEHR, INNOVATION UND TECHNOLOGIE,
Renngasse 5, A-1010 Wien
Dr. Rupert Pichler rupert.pichler@bmvit.gv.at

bm vit

BUNDESMINISTERIUM FÜR WIRTSCHAFT UND ARBEIT,
Stubenring 1, A-1010 Wien
Mag. Wolfgang Neurath Wolfgang.NEURATH@bmwa.gv.at

BMW
BUNDESMINISTERIUM FÜR
WIRTSCHAFT und ARBEIT

FFF-Forschungsförderungsfonds für die Gewerbliche Wirtschaft,
Kärntnerstraße 21-23, A-1015 Wien
Mag. Klaus Schnitzer klaus.schnitzer@fff.co.at

FFF

FORSCHUNGSFÖRDERUNGSFONDS FÜR DIE
GEWERBLICHE WIRTSCHAFT

FWF-Fonds zur Förderung der wissenschaftlichen Forschung,
Weyringergasse 35, A-1040 Wien
Dr. Rudolf Nowak novak@fwf.ac.at

FWF

Der Wissenschaftsfonds.

JOANNEUM RESEARCH, Institut für Technologie- und Regionalpolitik,
Wiedner Hauptstraße 76, A-1040 Wien
Mag. Wolfgang Polt wolfgang.polt@joanneum.at
Mag. Klaus Zinöcker klaus.zinoecker@joanneum.at

JOANNEUM

RESEARCH

KMU Forschung Austria,
Gusshausstrasse 8, A-1040 Wien
Dr. Sonja Sheikh s.sheikh@kmuforschung.ac.at

KMU FORSCHUNG AUSTRIA
Austrian Institute for SME Research



ARC Systems Research,
2444 Seibersdorf
Mag. Petra Wagner petra.wagner@arcs.ac.at


seibersdorf research
Ein Unternehmen der Austrian Research Centers.

TECHNOPOLIS,
Prinz Eugen Straße 80/12, A-1040 Wien
Mag. Leonhard Jörg leonhard.joerg@technopolis-group.com


TECHNOPOLIS

TECHNOLOGIE IMPULSE GESELLSCHAFT TiG,
Grillparzerstr. 7/8, A-1010 Wien
Dr. Dorothea Sturn dorothea.sturn@tig.or.at

TiG
Technologie Impulse Gesellschaft m.b.H.

ÖSTERREICHISCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG,
PF 91, A-1103 Wien
Mag. Hannes Leo hannes.leo@wifo.ac.at

WIFO
WIRTSCHAFTS- UND
WISSENSCHAFTLICHE
FORSCHUNG
AUSTRIAN INSTITUTE OF
ECONOMIC RESEARCH

WWTF – Wiener Wissenschafts-, Forschungs- und Technologiefonds,
Währinger Straße 3 / 15a, A-1090 Wien
Dr. Michael Stampfer michael.stampfer@wwtf.at

W W T F
Wiener Wissenschafts-, Forschungs- und Technologiefonds

ZIT – Zentrum für Innovation und Technologie GmbH,
Ebendorferstrasse 4/DG, A-1010 Wien
Robert Mayer-Unterholzner robert.mayer@zit.co.at

ZIT

zentrum für innovation und technologie gmbh

Rupert Pichler
Klaus Zinöcker

Preface

*“But when you come knocking at my door,
Fate seems to give my heart a twist.”*

Ted Koehler & Harold Arlen,
The Devil and the Blue Sea, 1931

In the most general of terms, peer review is the activity of having another researcher assessed the research you have done or intend to do. The peer attempts to identify the research's strengths and weaknesses, making stop-or-go decisions or suggests strategies for revising it. Peer review is, for sure, the most widespread evaluation method in the field of R&D; and (almost) irreplaceable when quality of projects and programs should be assessed.

Different forms of peer review are discussed in the continuing by John Rigby. Peer review exhibits a large capability of conflicts: peer review is often decision making about the individual work of a researcher or a research team. Although it is an old and well developed concept, it bears a wide range of possible shortcomings, that seem to be inherent in human nature: individual dependencies and relationships, ignorance and narrowness of the (scientific) horizon etc.

A recent Austrian study (Gerhard Streicher et al.: “Evaluation of FWF – Impact Study”, Joanneum Research, Vienna 2004) tried to highlight these possible shortcomings in the context of the Austrian Science Fund FWF. Beside the scientific ‘quality’ of a research proposal (which is to be assessed by external peers), there are other dimensions (like age,

profession etc.) which could play a role for funding decisions. Aiming at identifying parameters which influence the FWF’s decision on whether to accept or reject a certain proposal Streicher’s analysis shows that there seems to be no types of project which are either ‘natural losers’ or ‘natural winners’. Different submitted proposals find a ‘level’ playing field in relation to the probability of acceptance.

Peer review continues to be subject of methodological discussions around the globe, finding its manifestations in whole libraries on the pros and cons of peer review. However, it also continues to be an indispensable tool within sound evaluation procedures.

Need for improvement does not only or primarily arise from the problems inherent to Peer review. It is also a changing context of public funding instruments as well as a changing research culture leading to multi-disciplinary and multi-actor projects that often seem to be incompatible with the traditional realm of peer review. Assessing such projects is a challenge not only for peer review but for all techniques at hand.

Authors

Dr. Rupert Pichler
BMVIT - Federal Ministry for Transport,
Innovation and Technology
A-1010 Vienna, Renngasse 5
Phone: +43 1 53464-3205
rupert.pichler@bmvit.gv.at

Mag. Klaus Zinöcker
Joanneum Research &
Plattform Forschungs- und Technologieevaluierung
A-1040 Vienna, Wiedner Hauptstraße 76
Phone: +43 1 581 75 20-2827
klaus.zinoecker@joanneum.at
www.fteval.at

John Rigby

Making Decisions about Science and Technology - between the Devil and the Deep Blue Sea?

A brief reflection on Expert and Peer Review

INTRODUCTION

Anyone offering comments on peer review is more than unusually sensitive to the need to ensure that the views they advance are subject to thorough scrutiny by the relevant experts before claiming for them the status of statements of valid opinion or fact. This short article, while containing a number of personal observations about peer review, does not throw caution quite to the wind as much of what is covered has been explored in more depth in the EPUB project report on evaluation methodologies. However, since the report was written, a number of controversies have grown more intense and some changes that were over the horizon that will bring change to our current systems of peer and expert review have now come into view.

This article has four main sections. It begins with a look at origins of peer review as we encounter it in the 17th century, and note some of the features and pre-conditions for its successful operation which appear to have a remarkable historical resilience. It then looks at how review by experts has developed and how its value as a form of decision making on matters of uncertainty is evidenced by its mutation into a wide variety

of forms to meet an increasing range of purposes. Thirdly, I shall then look at review by experts today and the problems which face those who participate in some way with review by experts and who are affected by it. Finally, I shall note a few changes which are already on the horizon and what problems these raise.

ORIGINS OF PEER REVIEW – FEATURES AND PRE-CONDITIONS

Starting in the early modern period in the 1650s, peer review was developed as a procedure to assess the suitability of articles presented for publication to the Royal Society. In France, the Journal Des Scavans, composed almost entirely of notices of book publications was first published in 1665. The Philosophical Transactions of the Royal Society were published first in 1665. Other journals which followed were: Gionale dei litterati di Roma established in 1668 in Italy, the Miscellanea curiosa medico-physica, established in 1670 in Germany, the Acta medica et philosophica hafniensia established in 1673 in Denmark, the Collectanea medico-physica established in 1680 Holland, and the Acta eruditorum established in 1682 in Germany (Ornstein, 1963).

This is the originating form of decision making process that would be eventually generalized as expert review and which would have the following characteristics: a) asking those who are most qualified in a subject or area for their opinion in matters of fact and on matters of judgement that pertain to that area; b) intellectual exercise within a social process, an agonistic, according to

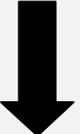
Ziman (1995) and involving a dialectic from which one might expect a more refined and sophisticated judgement, thereby earning for the process the reputation that it is indeed more than the sum of its parts; c) expert Review gives decisions, but not always “final decisions”; d) a dialectical, critical and investigative process rather than a fixed and formulaic approach may prevent distortions to its subjects; e) a technique suitable for many areas – if there are experts, a decision can be made: conversely, if there are no experts, by implication, any view is as valid as any other.

REVIEW BY EXPERTS – DEVELOPMENT OF DIFFERENT FORMS OR PRACTICES

During the last half millennia, and particularly during the last century, the forms of expert review have multiplied. As Table 1 shown below indicates, peer review is the most specialist type of expert review. Experts who are normally within the academic context are not paid for their work but give their time freely. Moving down the table generally

indicates a greater level of professionalization – i.e. a greater use of professional evaluators and a greater involvement of a wider range of experts from broader disciplinary backgrounds. The diagram is also intended to show an important fault line between traditional science and more modern, post 1950s science in their uses and types of peer review. Traditional academic science – Little Science (De Solla Price, 1963) or Republican Science (Fuller, 2000) employ Traditional Peer-Review (Canonical Academic Review) Direct Peer Review. By contrast, the Big Science (de Solla Price, 1963) Post-Academic Science (Ziman, 1995) & Liberalized Science (Fuller, 2000) models employ peer review within broader limits, opening up the process of decision making and reconceiving peer review as process involving stakeholders that can settle or decide questions of knowledge and also as dealing with the interests and priorities of different groups.

Table 1: The Forms of Peer Review

Sub-Type of Expert Review	Science	Level of Specialisation	Level of Professionalization
Traditional Peer-Review (Canonical Academic Review)	Academic Science Republican Science (Fuller, 2000) Post-Academic Science (Ziman, 1995) & Liberalized Science (Fuller, 2000)	Generally Increasing 	Generally Increasing 
Direct Peer Review			
Modified Direct Peer Review			
Pre-Emptive Peer Review			
Indirect Peer Review			
Merit Review (extended form of Peer Review)			
Ancillary Peer-Review			
Expert Panels/Peer Review			
Panel Review			
Professional Evaluators			
Extended Peer Communities	Post Normal Science (Funtowicz, Ravetz)	Specialisation non-relevant	Wider communities – anti-professional

THE PURPOSES OF PEER REVIEW

It is also worth considering in more detail the precise purposes and meanings which peer review takes across the range of functions outlined in Table 1. The Forms of Peer Review. There are, I believe, three main senses in which peer review can be understood: the review of claims to truth, the review of claims regarding the validity of methodology, and thirdly, the review, involving all interests deemed relevant, of the form of scientific knowledge that should be pursued and which serves best the interest of society.

1. Veracity & credibility of a claim to scientific knowledge

This is the traditional sense of peer review where peers argue over whether a claim to knowledge is justifiable. This is an ex post sense of review, in which the claim has been advanced, often on the basis of knowledge which is uncertain and difficult to interpret within existing conceptual frameworks.

2. Relevance of a test to affirm or to deny the truth of a claim to scientific knowledge

This second purpose differs from the first in that it concerns not whether a claim is valid or invalid, but the process, or test or experiment of what is likely to result in knowledge. This is review ex ante and may often refer to the scope, plausibility and scale of a test which has been proposed. The test is hypothetical in that there is, at the point in time when the review takes place, no direct evidence to judge of a claim to knowledge.

3. Science Serves Society

The third purpose to which peer review can be put is a wider application of peer review altogether involving a far broader definition of peer. In this sense of peer review, the aim is to draw upon wider forms of knowledge which have not traditionally been accepted within the process of determining the value of scientific research. This can be both an ex ante and an ex post form of review, although it is more likely as an ex ante form in which decisions are taken about what form of scientific or technological research is acceptable to society. This sense is close to a governance model of science; however, because the development and the consequences of technology are complex and knowledge often arises across domains, governance systems are not always able to consider and judge developments systematically.

The relationship of science to society in this third sense is to an extent a reversal from the first sense. While in the traditional senses of peer review, the experts or peers begin the process, in the second, peer review provides a check and decision making process over what knowledge is used by society.

Figure 1: Science and Society

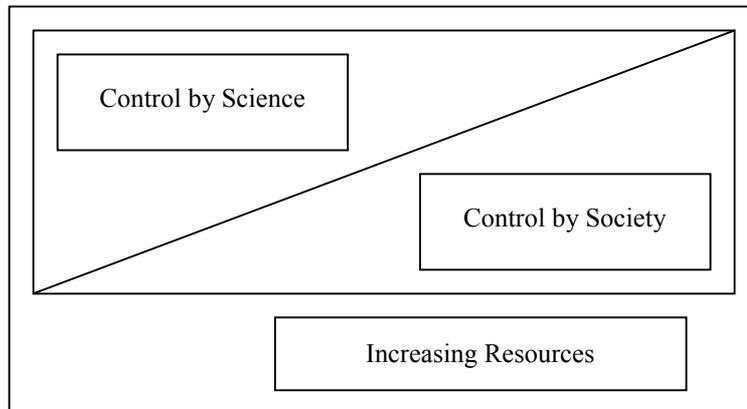
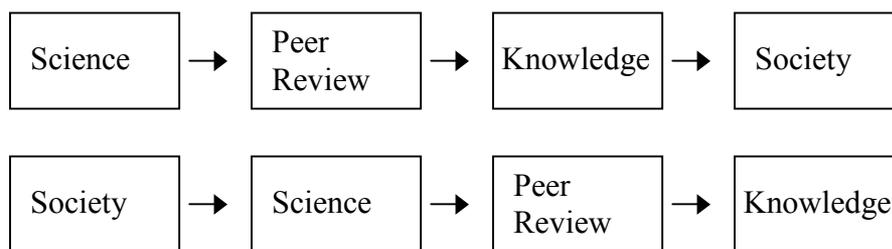


Figure 2: Science and Society – Two Views



APPLYING THE METHODS – PEERS AND PANELS

This next section examines the two forms of peer and panel review and considers some of the details of practical application. Panel and Peer Review methods are used prospectively, i.e. ex ante, and retrospectively, but the typical use of ex ante peer review is for the allocation of funding to research grant applications, while the most typical form of the Panel Review process is the ex-post evaluation of RTD programmes.

Panel Review is most likely to give reliable results where general broad brush pictures of the quality of research are carried out and

where evaluators and policy makers need a general picture of what impacts on socio-economic development have occurred. Panel review can also be thought of as a last resort in that it can be used where no other means of generating data or opinions are available. As panel review is a dialectical process, in its ideal form, it has the potential to provide new ways of looking at programmes and new insights into programme performance.

It is also worth noting that as panels are often the first to face an evaluation challenge in the either the ex ante or ex post sense, they can be the first to experience the need for new evaluation methods. As the EPUB report has

said, "Panels as evaluators may also contribute to the development of research structures and systems. In the case of the EUREKA Programme, a major contribution to the development of the Programme was the introduction of the continuous and systematic evaluation methods which were subsequently adopted for use in the evaluation of the Programme."

Pre-emptive peer review is the form of review in which the whole of the decision-making process is placed in the hands experts. Here, the experts therefore have the sole right and unchallenged right to make decisions. As this type of review removes the element of discretion from the appointees, the technique is not used widely, although the National Institutes of Health (NIH) in the United States have operated evaluation in this form occasionally (Bozeman and Melkers, 1993).

Ancillary Peer Review uses peer review along with other methods to decide a question of value or impact. Modified Direct Peer Review uses traditional peer review but peers are asked to look for broader impacts than at the narrow definition of the quality of publication outputs.

While Expert Review is often regarded as a specific research evaluation methodology, a Panel or Expert Review team is free, if its terms of reference allow, to carry out further studies which employ other research methodologies, such as surveys, questionnaires, benchmarking and case

studies. Expert Review, whatever its form it takes therefore, is more than simply a specific method, such as benchmarking or econometric analysis, but a system in which judgement is delegated to a set of independent or quasi independent agents and to choose, if they wish, discrete methods of data collection and analysis.

CURRENT PRACTICE AND FACING THE PROBLEMS OF PEER REVIEW

The text of the EPUB report covers in some detail but not exhaustively the strengths and limitations of peer review and shows what good practices for peer review have come into existence. It is not possible however, to be entirely comprehensive and exhaustive on these points of good and bad practice as many subtle variations to the process of peer review have been developed to meet various specific challenges thrown up by the need to control the costs of peer review, the need to include international peers or referees, and need to take account of the small size of some communities of researchers. There is, I think, no need to enumerate all of these here, but rather it would be better to examine some of the main issues and to explore a conceptual model which sees peer review as a process that has strengths and limitations that result from internal (endogenous) or design issues and from outside (exogenous) sources.

Table 2: Strengths and Weaknesses of Peer Review

Strengths	Limits
Perceived Independence	Scope of application is defined by availability of experts
Panels can employ methods which their specialist expertise suggests might be relevant, for example Scientometrics techniques such as bibliometrics, or surveys	Risk of geographical bias, self-interest
Broad evaluation of quality approach which specialists are unable cover	Programme Officers could manipulate evaluators
Inexpensive	Peer Review is subjective – it is best supplemented by technical – metric based measures
Can be continuous	Costs can be significant, although they may be born by their panellists and their organisations rather than the client
Generally flexible - can work off-line	Peer Review and variants shown to be conservative and leading to “institutionalised orthodoxy” (US Department of Health and Human Services, 1989, quoted in Bozeman and Melkers, 1993)
	Where research is likely to be non-public domain, the peers might be competitors. In this case peer review is not helpful.
	There may be a reluctance on the part of panellists to state their views on paper as peer review is normally non-anonymous (Grigson & Stokes, 1993)

PROBLEMS INSIDE

Endogenous difficulties with peer review arise at the levels of conceptual, representational and social interaction, all of which can undermine to some degree the decisions which a group of peers may reach by way of systematic or random bias. These internal difficulties will include problems with the identification of suitable peers, and this can be problematic for funders, whose knowledge of the subject area might not be as extensive as that possessed by the applicants. Limitations set on the geographical area from which peers might come can also give rise to bias in the judgement. Limitations can often be set for economic and efficiency reasons, and not in order deliberately to skew the findings of peer review.

Groups of peers need to comprise those with the most deeply rooted in theory and practice of a scientific area. However, such disciplinary specialists are often ill-equipped to decide how to treat new claims which challenge existing theory and practice, and which cut across disciplines as interdisciplinary and transdisciplinary do.

Observers of peer review also note that peer review is a social process involving individuals, and, to some extent, the institutions from which they come. Within academic peer review, therefore certain peers may exert greater control over the process than applicants often believe is acceptable, taking a role in the interpretation of findings that extends beyond the function of judgement either for or against a view presented to them (Frey, 2003). Blinding the

review process does not always result in fairness as closed groups tend to be limited in their views.

It should also be borne in mind that the peers themselves are often in competition with those whose work they judge. In consequence, there are often incentives on the part of peers as reviewers either ex ante or ex post not to give complete information about the work they judge (Schmidt, 2003) for fear of losing priority or for losing an advantage. Indeed, pursuing this gaming analogy, peers could suggest false trails for applicants to follow; but this is a long way away from any desirable form of peer review.

PROBLEMS OF CONTEXT

Turning to exogenous factors, when peer review is manipulated from the outside, there appear to be two main forms. On the one hand, powerful groups can influence the selection of peers. Governments and funding agencies can do this if they wish, although they are likely, in the longer term to harm their reputation and those of the honest peers they do include.

There are also more subtle methods that involve the creation of spurious literatures or political pressure that seeks to sway the judgement of honest and unbiased peers. Barnett (2003) has reported that institutions wishing to influence peer review have been found to be manipulating the medical literature by employing medical practitioners to promote a particular line of research behind which lies a form of treatment in

which particular pharmaceutical companies have an interest.

ON THE HORIZON – SOME COMING DEVELOPMENTS

Current interest in peer review stems from some of the problems raised above which appear now to be more salient and pressing than hitherto. But there is also interest because of developments in political institutions and in technology which can be used in conjunction with peer review.

Regarding the issues raised in the earlier part of this article, there is now an increasing amount of concern with and scepticism about the benefits of science and these stem from major health crises such as BSE in the UK, the development of genetically modified organisms and the appearance of technologies that could support human reproductive cloning. Our scientific knowledge now poses questions for society of moral, legal and political nature and of a fundamental kind.

In regard to the institutional and technical matters, two issues are of prime importance: how should research be organized within the European Research area; and what role can bibliometrics analysis play in the judgement of research outputs and allocation of funds to research.

The proposals for a European Research Council which are currently under discussion suggest a wholly new Europe wide research council through which funds for academic research will be allocated. There are great expectations that such a system will ensure scientific excellence at a European level,

helping the European Union to attain even high levels of scientific achievement than in the past through combining the skills and capacities at European and therefore larger level. At such a larger level, research funding reaches a critical mass that can support the investigation of areas of science that might not be possible to pursue within smaller systems. The new system will be based on peer review. The EURAB recommendations envisage the following:

“An ERC should be founded on autonomous scientific decision making based on the peer-review process but accountable to a representative (political) governance structure. Such a political structure has to be at arms length from the science decisions. However, this structure can also serve as a route for the two-way transfer of ideas between the political and the scientific ‘worlds’. One model in which Science or Research Ministers form a ‘trusteeship’ council may provide sufficient accountability without undue involvement in the autonomous scientific decision making.”

The challenge here for the direction and orientation of science is how to establish the interface between science and governance, particularly as governance at a European level has proved such a problematic issue. The question of how peers will be directed and the balance of top down and bottom up funding will remain the fundamental issue for the ERC that it is for all research councils. Furthermore, and regarding the issues of scientific capabilities – the input side - there will be tension as the effects of investigator-led science make more stark national differences in research capacity. This will lead to calls for protection for particular

(national) thematic and organisational resources.

Aside from this major institutional change taking place, in the area of technology of bibliometrics, innovation in information technology is giving both scientists and research funders more information about the impacts of funding. The new techniques are providing benchmarks to make comparisons between individual scientists, the institutions in which they work and the decisions of the funding agencies which supported them.

This allows raw measures of reputation and efficiency to be calculated, and for funders, this is a particularly attractive technique with which to assess the consequences of their actions. However, there are two major problems: firstly, benchmarking input and outputs measures across different national science systems and different subject areas is a non-trivial problem. National differences are often profound, and the resulting indicators can be wholly misleading. Secondly, while bibliometrics data already constitutes a derived measure from a wider peer review process that involves scientists reading each others work and citing it in what they subsequently publish, such measures should not be used to decide on scientific priorities or to decide the merits of a particular funding application. This should always be done on the individual strength of the scientific case with trust in the judgement of those most suitably qualified.

PEER REVIEW – AN AFTERTHOUGHT

A survey of a literature in any subject will in due course cover a wide range of references some of which will present extreme views on aspects of the debate. Anyone examining the interesting subject dealt with in this article will be unlikely to encounter the opinions of what is knowledge belonging to the English clergyman, Benjamin Jowett, and sometime Master of the Oxford college, Balliol, but they are included here for the sake of amusement. Jowett's view of how we know that something is knowledge represents one of these extreme positions; but there is little reason to take notice of him now.

*Here I come, my name is Jowett
All there is to know, I know it
What I don't know, is not knowledge
I am the Master of this College*

Benjamin Jowett (1817-1893)

References

The references below are in two sections, general references and references concerning EU RTD specifically. The first section has been expanded from the EPUB report to include material used for a presentation made in Vienna in 2003 at the Austrian Ministry by the author.

Barnett, A. (2003) "Revealed: How drug "hoodwink" medical journals" *The Observer*, 7 December 2003, page 15.

Budd, John M. "Silencing Scientists and Scholars in Other Fields: Power, Paradigm Controls, Peer Review, and Scholarly Communication. (Book Reviews)" (book review) *Libraries & Culture*, Spring 2000 v35 i2 p.380(2).

EURAB WORKING GROUP 4 (2003) *The European Research Council (ERC) - A*

Possible Implementation Model

- Frey, B. S. (2003) *Publishing as prostitution? – Choosing between one's own ideas and academic success* *Public Choice* 116: pages 205-223.
- Guston, D. (2000) "Expanding the Role of Peer Review Processes in the United States", paper presented to the US-European Policy Workshop on Learning from Science and Technology Policy Evaluation, Bad Herrenalb, Germany.
- Chalk, R. (2003) *Gordon Moran. Silencing Scientists and Scholars in Other Fields: Power, Paradigm Controls, Peer Review, and Scholarly Communication. (Book Review)* *Isis*, Sept 2003 v94 i3 p552(3).
- Campbell, P. (2001) "Opinion: Introducing a new policy for authors of research papers in *Nature* and *Nature* journals" *Nature* 412, 751 (2001).
- Economist* (2003) "Perishing Publishing" November 15th, page 105.
- Funtowicz S., and Ravetz J.R. (1993) *Science for the Post-Normal Age. Futures*, 25:735-755.
- Godlee, F. & Jefferson, T. (1999) *Peer Review in the Health Sciences*, London: BMJ Books.
- Grigson, D. & Stokes, T. (1993) "Use of peer review to evaluate research outcomes" *Research Evaluation*, Vol. 3, No. 3, pages 173-177.
- Holton, Gerald (1978) *The Scientific Imagination: Case Studies*, Cambridge: CUP.
- Hyland, K. (2000) "Disciplinary Discourses: Social Interactions in Academic Writing", Harlow: Longman Publishing.
- Lievrouw, Leah, A. (2000) *Silencing Scientists and Scholars in Other Fields: Power, Paradigm Controls, Peer Review, and Scholarly Communication. (Review)* (book review) *Library Quarterly*, Jan 2000 v70 i1 p163.
- Moran, G. (1998) *Silencing Scientists and Scholars in Other Fields: Power, Paradigm Controls and Peer Review, and Scholarly*

Communication Greenwich, Connecticut: Ablex Publishing.

- Ravetz, J.R. (1999) *What is Post-Normal Science Futures 31 647–653*
- Ravetz, J. & Funtowicz S. (1999) “ *Post-Normal Science—an insight now maturing*” *Futures 31 (1999) 641–646*
- Roy R (1984) *Alternatives To Review By Peers - A Contribution To The Theory Of Scientific Choice Minerva 22 (3-4): 316-328.*
- Roy R (1985) *Funding Science - The Real Defects Of Peer-Review And An Alternative To It Science Technology & Human Values (52): 73-81.*
- Sackett, D., and Oxman, A. (2003) *HARLOT plc: an amalgamation of the world's two oldest professions BMJ 2003;327:1442-1445 (20 December)*
- Schmidt, A. (2003) – *Discussion with the author.*
- Weinberg, A. (1966) *Reflections on Big Science, Cambridge: MIT Press.*
- Ziman, J. (1995) *Of One Mind: The Collectivization of Science, Woodbury, New York: AIP Press.*
- Ziman, J. (1998) *Silencing Scientists and Scholars in Other Fields: Power, Paradigm Controls, Peer Review and Scholarly Communication. (Review)_(book reviews) Nature, Oct 29, 1998 p856(2).*
- Ziman, J. (2000) *Real Science, What it is, and what it means, Cambridge: Cambridge University Press.*
- considerations”, Research Evaluation, Vol. 21, No. 6, pages 405-413*
- Bobe, B. & Viala, H. (1997) “*Panels of experts of the European Commission*” in Callon, M., Laredo, P. & Mustar, P. (eds.) *The Strategic Management of Research and Technology – Evaluation of Programmes”, Economica: Paris*
- Bozeman, B. & Melkers, J. (eds.) (1993) *Evaluating R&D Impacts: Methods and Practice, Kluwer: Boston*
- Campbell, P. (2001) “*Opinion Introducing a new policy for authors of research papers in Nature and Nature journals*” *Nature 412, 751 (2001)*
- Massimo, L. (1997) “*Lessons from the experience of the European Commission*” in Callon, M., Laredo, P. & Mustar (eds.) *The Strategic Management of Research and Technology – Evaluation of Programmes”, Economica: Paris*
- Porter, A. & Rossini, F. (1985) *Peer Review of Interdisciplinary Proposals” Science, Technology and Human Values, Vol. 10, No. 3, pages 34-42*
- The Economist “Truth or Consequences”, September 15 2001*

Author

Dr. John Rigby
 Research Associate
 PREST - Policy Research in Engineering Science and Technology
 Manchester University
 Oxford Road, Manchester, M13 9PL
 Phone: +44 (0)161 275 5928/5921
 Fax: +44 (0)161 273 1123
John.Rigby@man.ac.uk
<http://les.man.ac.uk/PREST>

Operational Guidelines to Evaluation

The following are works on EU peer and especially panel review. The EPUB review notes that there is only a small amount of work on the evaluation of panel review within this context.

Barker, K. (1994) “*R&D Evaluation – Strengthening the impact of R&D evaluation on policy making: methodological and organisation*

Michael Stampfer

Some Developments in Peer Review

INTRODUCTION

When it comes to appraise and evaluate the scientific quality of a funding proposal, a research group, an institution or a paper, peer review is an indispensable part of this quality assurance process. From written reviews to extensive site visits, from anonymous statements to discussions with fully developed feedback circles, a broad range of forms and methods can be observed. The core is always the same: A sufficient number of impartial experts from a given field are asked – normally by decision makers like funding bodies, managers or editors – for their opinion and judgement about a certain plan or achievement. The core qualification of the peers is therefore their internationally renowned command of the specific scientific field or sub-field in question, usually not less, but also not more.

In Austria peer review is common practice in nearly all funding institutions dealing with scientific quality. The Austrian Science Fund (FWF, the Austrian research council) uses foreign peers for all proposal evaluations, so do other smaller funds and specific programmes. In addition to written statements, juries and site visits have become popular in the last years. Mirroring the overall situation of RTDI evaluations in Austria, peer review is strong in ex ante

evaluations of projects, weaker but growing in interim and ex post project evaluation and in programme evaluation and considerably weak in any form of institutional evaluation.

This contribution first gives a short overview of requirements, strengths and weaknesses of this approach, also including some remarks Arnold Schmidt stated as one of the speakers in the Platform's peer review workshop. In a second part peer review is discussed as a valuable, even indispensable evaluation element of complex funding programmes aiming at applied and / or interdisciplinary research. In this kind of programmes or comparable funding initiatives "pure" peer review is seen as insufficient and has to be extended by other kinds of expertise and in some cases by other methodological approaches.

SOME MINIMUM REQUIREMENTS, STRENGTHS AND WEAKNESSES OF PEER REVIEW

There is a common understanding that peer review is an indispensable instrument of quality appraisal, namely in all forms of ex ante evaluations. This is also normally the starting point for all kinds of "but ..." - arguments, ranging from technical criticism to more fundamental issues including the claim of systematic bias.

There seems to be a number of minimum requirements for peer review. They are of course not static but depend on the size, form and properties of the subject under review. The first point is proportionality; there is no such as "one size fits all" approach. Nevertheless even small projects and initiatives should generally be reviewed by more than one outside expert: Bias and error

are everywhere, a second or third opinion helps to minimize this kind of problems. The second and most important point is impartiality: The reviewer shall know the field well, but should not have any specific interests in the subject reviewed. This raises tricky questions about the necessary closeness vs. distance and it imposes duties on the organising (funding / editing) body in e.g. checking co-publications or asking the reviewed party to name certain persons as potentially biased. A corresponding duty is on the reviewer's side: He / she has to disclose links, conflicting roles and interests. On a macro level a number of namely smaller countries rely practically exclusively on foreigners as reviewers. A third, related issue regards clear roles and a conscious handling of these roles. Programme managers and funding council committee members for example are important gate-keepers, there are some cases where it is even better to ask foreign institutions to nominate the peers. A fourth minimum requirement regards clear and transparent procedures regarding inter alia timetables, criteria, anonymity, feedback and impacts.

The strengths of peer review are quite obvious: Again the impartial judgement of experts with the same background is most important for the ex ante appraisal and the evaluation of proposals. This goes along with deep understanding of thematic and methodological issues. The strongest argument for peer review is to be found in the combination of these two strengths: It is needed and nobody does it better. Given the specialisation in the scientific world no in-house appraisal, no mechanized / purely quantitative indicator system and no external

know how apart from scientific peers will generally work. As a by product of peer review, funding organisations (or policy makers or journal editors) can build up international expert networks. By the way foreign peers are deliberately seen as one building element for research areas like ERA as long as comparable standards, procedures and criteria are used by different funding organisations for comparable kinds of activities. In this respect peers can help to build – in the case of ERA – good practices by cross-national review procedures. A final and often underestimated strength regards the ability of peers also to give their judgement on organisational and managerial issues when evaluating research programmes, centres and institutions, a strength that – see below – needs to be complemented.

The weaknesses and shortcomings of peer review as an evaluation instrument have been strongly discussed in the last years. Apart from some criticism from the corner of quantification fundamentalists most of the objections are centred around the old boys network argument. There are indeed many indications that peer review favours a) men compared to women and b) established researchers compared to newcomers, younger researchers and outsiders to the system. The “Newcomer”-argument is also often used to discuss the suspicion that peer review favours mainstream careers, proposals and projects compared with new and namely interdisciplinary approaches. The “Guardians of Science” are further criticised for having too many interests of their own. All scientists in a field are to some extent

competitors in a world wide market for fame, positions, strategic advantages, money, discovery and it's dissemination. So funding organisations have to take into account that reviewers could be tempted to use the knowledge they gained in the evaluation process for their own purposes up to plagiarism or to help mediocre researchers / downgrade very good ones along their alliances and self-interests. National borders are not a natural boundary in this respect.

These arguments, most of them brought forward by Arnold Schmidt as speaker in the Platform workshop, need to be taken into account by funding organisations. In all cases ways to ease the problems described exist. Meaningful alternatives are lacking and there is clearly no argument to turn peer review down.

As mentioned some of the critical points are centred around the issues of new fields, interdisciplinary and / or application oriented research. Here peer review is one important element of all evaluation steps but needs some adaptations and combinations with other methods and different know-how.

NEW CHALLENGES AND FIELDS FOR PEER REVIEW

As stated above peer review depends largely on the subject we are dealing with. This is true within the scientific paradigm, i.e. it matters whether we talk about a publication, a career step, a project, a research programme or a scientific institution. When it comes to RTDI policy instruments, where scientific quality is one important element within a broader array of criteria, peer review

in many cases has both to be adapted and complemented.

Typical policy instruments are RTDI funding programmes aiming at interdisciplinary research and / or science – industry cooperation. Both have their rationale in innovation systems, i.e. the need to better link up different actors, and in new scientific developments stemming from interfaces between disciplines or feedbacks from technological developments. Competence centre funding programmes for example have appeared over the last decade in a number of countries, from the U.S., Canada and Australia to European countries like Sweden, Germany, Austria or Hungary. Within such programmes scientific institutions team up with a number of firms to run a multi-annual, precompetitive but cooperative research programme. Issues of scientific quality are as important as industrial relevance and managerial / organisational questions. Another example are programmes to fund interdisciplinary projects, be it converging technologies with a longer time horizon or rather application or solution oriented interdisciplinary initiatives. Here the appraisal of quality of individual fields and sub-disciplines has to be complemented by a synoptic view.

So if policy makers plan such programmes and design evaluation systems (i.e. ex ante-, ex post-evaluation and monitoring on project and programme level), they have to take these dimensions into account. They shall ask themselves what has to be found out at which stage by whom and integration of peer review in their evaluation system is one of the most prominent questions.

For the ex ante-, interim- and ex post-evaluation of larger projects, centres and initiatives extended peer review has become a method of choice. Though individually different, such peer centred procedures generally combine specific discipline-oriented know-how delivered by typical peers with abilities of a different kind: This can be expertise on industrial RTDI or organisational and managerial know-how in the case of ex ante-evaluations. In interim evaluations also organisational expertise matters, finally the skills of professional evaluators. Often senior scientists from different fields – but with a track record in research organisation – can play a prominent role in such procedures, which is also true for people with a strong professional research management (and funding) expertise.

Such forms of extended peer groups need interaction processes beyond written feedback loops. This is one reason why site visits, panels, evaluation groups and juries have become not only popular but indispensable for programme funding organisations. Such mixed groups with a strong chairperson can come forward with recommendations that are both deep (evaluation of individual people, approaches and disciplines) and broad (evaluation of links, interdisciplinary approaches and structural recommendations). The input of the “discipline” peers can be twofold: Aside from strict scientific matters they often provide most valuable comments on structural and organisational issues – but it is not the rule and it needs interaction with people in a peer group whose prime responsibility are such structural and organisational questions.

On the programme level we also see more such extended peer groups. In ex ante evaluations we find juries with mixed compositions, again ranging from scientists coming from the disciplines funded to industry people and funding specialists. Often there are combinations between two steps: A classical peer review for each project and in a second step a mixed jury with a broader view for the funding recommendations where such a jury can use the individual review as a kind of objective basis. When evaluating interdisciplinary or centre programmes in a later step a similar approach can be taken in the form of mixed evaluation panels including peers, which has become a kind of good practice.

For all levels not only these specific features but also the “normal” pitfalls of peer review have to be taken into account by policy makers and evaluators. Though in a mixed group we generally find a higher degree of social control and a broader agenda, such advantages can and should be furthered: The terms of reference should include explicitly issues beyond discipline-oriented appraisal. The mix of the jury / evaluation group should contain also a sufficient number of women and of younger people. There should be control mechanisms in place to prevent hidden agendas or bias. For a funding organisation it is the art of shaking a good cocktail while staying sober at the same time.

Finally peer review, namely the interim and ex post evaluation of larger centres, programmes and other initiatives of considerable size and duration can be combined with other forms of evaluation

methods. Here quantitative approaches like impact analysis or forms of social network analysis are nice to combine with the qualitative expertise of peers. Through this process multi-faceted results can be obtained, often with interesting cross-fertilisations. Complexity and evaluation budgets though limit the use of such multi-method approaches.

References

- Daniel, H.-D., *Guardians of Science. Fairness and Reliability of Peer Review*, Weinheim (1993)
- Joanneum Research / IPTS (eds.), *EPUB – RTD Evaluation Toolbox. Socio-Economic Evaluation of Public RTD policies* (2002)

TIG et al. (eds.) RoadMAP. Good Practices for the management of Multi Actors and Multi Measures Programmes (MAPs) in RTDI policy

Wenneras, C. / Wold, A., Nepotism and sexism in peer-review. Nature 387 (1997), 341 ff (auf deutsch dies.: Vetternwirtschaft und Sexismus im Gutachterwesen, in Kraus, B., Wissenschaftskultur und Geschlechterordnung, Frankfurt – New York 2000)

Author

Dr. Michael Stampfer
Managing Director
Vienna Science and Technology Fund
A-1090 Vienna, Waehringer Straße 3
Phone: +43-1-402 31 4
Fax +43-1-402 31 43-20
michael.stampfer@wwtf.at
www.wwtf.at

Der Newsletter der Plattform Forschungs- und Technologieevaluierung GesbR ist ein unregelmäßig erscheinendes offenes Forum zur Diskussion methodischer und inhaltlicher Evaluierungsfragen in der Forschungs- und Technologiepolitik.

© Wien 2004

ISSN: 1726-6629

Herausgabe und Versand:

Dr. Rupert Pichler, Dr. Michael Stampfer, Mag. Klaus Zinöcker
Plattform Forschungs- und Technologieevaluierung GesbR,
A-1040 Wien, Wiedner Hauptstraße 76, office@fteval.at

Für den Inhalt dieser Ausgabe verantwortlich:

Plattform Forschungs- und Technologieevaluierung GesbR
Mag. Klaus Zinöcker



PLATTFORM fteval
Forschungs- und Technologieevaluierung

PLATTFORM FORSCHUNGS- UND TECHNOLOGIEEVALUIERUNG

Die Plattform Forschungs- und Technologieevaluierung GesbR ist eine Initiative des österreichischen Bundesministeriums für Verkehr, Innovation und Technologie, des Bundesministeriums für Bildung, Wissenschaft und Kultur, des Bundesministeriums für Wirtschaft und Arbeit, Joanneum Research, WIFO, Technopolis, WWTF, ARC Seibersdorf research, der KMU Forschung Austria, der Technologie Impulse Gesellschaft, der Zentrum für Innovation und Technologie GmbH, sowie den Fonds FFF und FWF.

Im Rahmen der Plattform werden Themenstellungen zur Forschungs- und Technologieevaluierung erarbeitet und – z.T. unter Einbeziehung namhafter ExpertInnen – in einem Fachkreis diskutiert. Der Newsletter beinhaltet Fachbeiträge zu Fragen der forschungs- und technologiepolitischen Evaluierung. Die Herausgabe erfolgt in zeitlicher als auch inhaltlicher Abstimmung mit Plattform-Veranstaltungen, um die Synergiewirkungen eines breiten Austauschforums zu nutzen.

