

Peer review

Its present and future state

Prague, 12-13 October 2006



European Science Foundation (ESF)

The European Science Foundation (ESF) was established in 1974 to create a common European platform for cross-border cooperation in all aspects of scientific research.

With its emphasis on a multidisciplinary and pan-European approach, the Foundation provides the leadership necessary to open new frontiers in European science.

Its activities include providing science policy advice (Science Strategy); stimulating co-operation between researchers and organisations to explore new directions (Science Synergy); and the administration of externally funded programmes (Science Management). These take place in the following areas: Physical and engineering sciences; Medical sciences; Life, earth and environmental sciences; Humanities; Social sciences; Polar; Marine; Space sciences; Radio astronomy frequencies; Nuclear physics. Headquartered in Strasbourg with offices in Brussels, the ESF's membership comprises 75 national funding agencies, research performing organisations and academies from 30 European nations.

The Foundation's independence allows the ESF to objectively represent the priorities of all these members.

Czech Science Foundation (GAČR)

The Czech Science Foundation (GAČR) was established in 1993 as an independent institution.

The basic aims are:

- to provide financial support for excellent research projects and at the same time to audit effective use of the financial means;
- to promote high-level research through long-term funding, based on peer review evaluation of submitted proposals, science-policy expertise and global cooperation;
- to raise the public understanding of science and to enhance the esteem and social status of scientific research;
- to develop high-quality research environments and to improve the scientific career opportunities;
- to support multi- and interdisciplinary research projects and communications;
- to represent Czech science among the international research organizations and in national and international scientific bodies;
- to cooperate and support international scientific co-operation of research projects through agreements with research councils all over the world.

European Heads of Research Councils (EuroHORCs)

EuroHORCs is the association of the Heads of public national research and research funding organisations in Europe. It was established in 1992 as an informal association of national research councils and analogous public non-university research organisations of the EU Member States. The last few years EuroHORCs has become an active player in the field of European research policy by promoting and enhancing inter-council cooperation and serving, amongst others, as advisory body for the European commission.

EuroHORCs seek to enhance the role of the national research and research funding organisations in Europe through creating a platform for discussion, initiating joint activities and strengthening their influence on European research policy.

Contents

Foreword from the Programme Committee	4
1. Introduction	5
2. Plenary Sessions	
Peer Review: Current Practices, Challenges and New Directions	
2.1. Peer Review in pan-European Research Funding Schemes	7
2.2. Peer Review in National Funding Agencies	8
2.3. Peer Review in Scientific Publishing	9
3. Parallel Sessions	
Peer Review: A Look From Different Angles	
3.1. Assessment and Selection of Research Proposals: Challenges and Procedures in National Research Agencies [Sessions A1 and B1]	11
3.2. Assessment and Selection of Research Proposals in International Organisations and Transnational Funding Schemes [Sessions A2 and B2]	14
3.3. Peer Review in the Evaluation of Research Institutions [Sessions A3 and B3]	18
3.4. Science Policy Context for Excellent Peer Review [Session C1]	21
3.5. Language, Metrics, Impact: Cultures of Peer Review in the Humanities [Session C2]	23
4. Concluding Session	26
Appendices	
1. Conference Programme	30
2. Conference Participants	32

Foreword from the Programme Committee

On 12 and 13 October 2006, the Czech Science Foundation, EuroHORCs and ESF organised an international conference on “Peer Review – its Present and Future State”. The conference was hosted by the Czech Science Foundation in Prague.

High level representatives of major research-funding and research-performing organisations were invited to present their approaches to the peer review process in their respective institutions. The enormous response to this invitation is an indication of the great interest in this topic. It strengthened our shared belief that presenting and discussing the practices of peer review with an international audience offers an opportunity to learn from each other. The conference, with more than 30 contributions, was organised in two plenary and eight parallel sessions. The conference programme is reproduced in Appendix 1.

The conference was attended by more than 150 delegates from European institutions as well as institutions from the United States, China, Japan and South Korea (See List of Participants, Appendix 2).

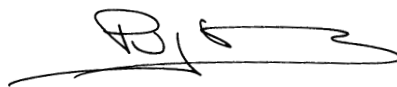
The Programme Committee would like to thank all conference speakers and participants for their active contribution, which led to lively and interesting discussions from a great variety of perspectives.



From the left: P. Nijkamp, J. Syka, B. Andersson



Bertil Andersson
Chief Executive of ESF



Peter Nijkamp
President of EuroHORCs



Josef Syka
President of GAČR

1. Introduction

The unprecedented pace of change in science systems in general and in the research landscape in most fields brings about new challenges to which the peer review system must quickly adapt. Continuous dialogue among concerned organisations is therefore needed. This dialogue provides an occasion for sharing experiences and spreading good practices and offers an opportunity for learning from each other how arising problems are being addressed, which practices work and which problem areas can be tackled jointly.

In this context the European Heads of Research Councils (EuroHORCs), the European Science Foundation (ESF) and the Czech Science Foundation (Grantová Agentura České Republiky, GAČR) joined forces to organise an international conference on peer review. The aim of the conference was to analyse contemporary trends in the evaluation of research, to examine how the peer review process is understood and performed, and to consider its future modifications in response to the requirements of research in the 21st century.

As a general focus, the theme of the conference was organised along three questions:

- Is peer review in the present form able to identify the best and most innovative frontier science and how might it be improved?
- What is the best way to harmonise the peer review process and how can new methods and IT tools contribute to it?
- What are the major societal, cultural and ethical challenges of future peer review processes and how could they be incorporated?

In some sessions, however, these questions led to the identification of additional issues that required discussion. There was room for this flexibility during the conference.

This report is a summary of the presentations and discussed issues. Most of the slides presented are available at the conference homepage (<http://www.pragueforscience.cz>). The succinct summaries given here are to guide readers to the presentations of their particular interest on the homepage.

We hope that this conference and this report will be useful in further discussions about improvements to the peer review system and will stimulate further exchange of information and experience among research organisations.

Programme Committee

Bertil Andersson, Chief Executive of ESF
Peter Nijkamp, President of EuroHORCs
Josef Syka, President of GAČR

Organising Team

Daphne den Hollander, International Relations, NWO
Nina Kancewicz-Hoffman, Science Officer in the CEO Unit, ESF
John Marks, Director of Science and Strategy, ESF
Veronika Palečková, Department of International Relations, GAČR
Katerina Sobotková, Department of International Relations, GAČR

2. Plenary Sessions

Peer Review: Current Practices, Challenges and New Directions



Opening of the conference (from the left V. Pačes, President of the Academy of Sciences of the Czech Republic, P. Nijkamp, President of EuroHORCs, M. Kopicová, Minister of Education, Youth and Sports of the Czech Republic, I. Halliday, President of ESF and J. Syka, President of the Czech Science Foundation)

Plenary Session I

Chair: **Ian Halliday**, *European Science Foundation*

Graham Stroud, *European Commission*, Evaluation rules and procedures under the EU Framework Programme for Research and Technological Development: the current state of play and future developments

Toni Scarpa, *National Institutes of Health*, NIH peer review: challenges and opportunities

Philip Campbell, *Nature*, Pressures on peer review at *Nature* journals

Kurt Mehlhorn, *Max Planck Society*, Peer review in the electronic age

Stefan Hornbostel, *Institute for Research Information and Quality Assurance*, Peer review – in the core healthy or chronically ill?

John O'Reilly, *Engineering and Physical Sciences Research Council*, Risk, adventure and the tyranny of peer review

Plenary Session II

Chair: **Josef Syka**, *Czech Science Foundation*

Arden L. Bement, Jr. *US National Science Foundation*, The reason and reach of merit review at the US National Science Foundation

Bengt Nordén, *Chalmers University of Technology*, Peer review and young research careers: EURYI scheme

Daniel Estève, *European Research Council*, Peer review system of the European Research Council

Motoyuki Ono, *Japan Society for the Promotion of Science*: JSPS's application review system – present and future

Stephen J. Simpson, *Science Magazine*, Exploring the limits of peer review in science publishing

2.1. Peer Review in pan-European Research Funding Schemes

In recent years, Europe has witnessed the development of new research funding programmes to improve the collaboration of researchers across national borders. Most of those competitive grant schemes rely on peer review mechanisms for the selection of proposals. In the plenary sessions of the conference, approaches used in three European research-funding schemes were presented: the EU Framework Programme for Research and Technological Development, the European Young Investigator (EURYI) Awards of EuroHORCS and ESF and the Starting Independent Researcher Grant scheme of the European Research Council.

Dr Stroud from the European Commission (EC) described the process of proposal selection for EU Framework Programme and discussed the improvements to be made for the Seventh Framework Programme (FP7) starting in 2007.

After proposal submission which, depending on the programme, may be in the form of a pre-proposal, an eligibility check is made according to requirements specified in the call for proposals. The proposals are then reviewed by external evaluators selected from a large pool of possible referees. Individuals listed in this pool are either proposed by institutions or by themselves in response to calls to register as potential evaluators. The EC may also select evaluators not listed in this pool. The evaluators sign a declaration of confidentiality and a conflict of interest form. Their names are published after the evaluation.

The evaluation criteria are pre-defined and may be differently weighted according to research area and the aim of the research programme. In FP6, these criteria included ethical considerations and gender issues in addition to science and societal issues.

Each proposal is reviewed by three or more evaluators who give marks and comments, on the basis of which a consensus is sought. A panel meeting compares the consensus reports and makes suggestions on the order of priority. During this process, hearings with grant applicants may be convened. The European Commission sends the evaluators' evaluation summary reports to grant applicants, draws up the final ranking list and takes the final funding decisions based on the advice of the experts.

For the future, the EC aims to further improve current procedures while maintaining continuity with current practices in the Framework Programme which proved to be efficient.

Improvements include the increasing use of IT tools to enable reviewers to have remote access to the proposals. Efforts to include high quality researchers

in the review pool will continue, in particular through using remote evaluation (which minimises the length of time busy scientists spend in Brussels). The criteria, adapted to each instrument, will be divided into three distinct sets: the scientific (and technological) quality of the proposal, the likely impact and the quality of implementation.

Professor Nordén presented the selection procedures used in the European Young Investigator (EURYI) Awards. This funding scheme was created in 2003 by the EuroHORCs in collaboration with ESF. Currently 15 countries participate in the scheme; to date 75 applicants have been selected for the award to create an independent research group. The selection of the awardees is organised in two steps: the first selection step is handled by the national research councils and the second step is undertaken by international panels established by ESF. Approximately 120 proposals enter the second stage. Out of which 25 are selected.

Professor Nordén focused on the second stage in which six broad disciplinary panels make the assessment. In establishing the panels (with about eight members each) scientific record is the main criterion and a balance of gender, expertise and geography is taken into account. All panel members read and individually score the applications. Grant applications deviating from the average are discussed and a preliminary list is established. Starting from the top of the tentative list, all applications are discussed, each with a selected spokesperson. This results in a rearranged list with decisions about which applicants to invite for an interview. About 60 candidates are interviewed. The interviews are seen as important because they give insights into the true potential and creativity of the candidate.

Dr Estève, a Vice-Chairman of the Scientific Council of the European Research Council (ERC), presented the peer review process to be used when the ERC starts its activities in 2007. The launch strategy of the ERC foresees two funding schemes on a bottom-up basis: the ERC Starting Independent Researcher Grant scheme (ERC Starting Grant) and ERC Advanced Investigator Researcher Grant scheme (ERC Advanced Grant). The ERC Starting Grant will support researchers at the start of establishing their first independent research team in all areas of research. It is expected that between 200 and 300 grants per year will be awarded.

The ERC developed a peer review mechanism in which the review of the proposals will be made by disciplinary panels (with about 12 members each) and will be done in two stages: First selection by the panel of twice the number of grants to be distributed and second panel evaluation based on reviews and interviews. The criteria are: the excellence of the

2. Plenary Sessions

project submitted and the potential of the applicants (research excellence, achievements, and publications). To identify outstanding researchers to serve on the 20 panels, major scientific organisations in Europe were approached to nominate candidates. Using these nomination lists and other sources, the scientific council of the ERC has identified a large pool of names (about 500 names) suitable for the constitution of balanced panels, inclusive in all respects.

2.2. Peer Review in National Funding Agencies

The plenary sessions also included presentations on the peer review process in two US and one Japanese organisations: The National Institutes of Health (NIH), the National Science Foundation (NSF) and the Japan Society for the Promotion of Science (JSPS). Case studies of two European institutions complemented the picture on challenges faced by peer review systems in national research funding agencies.

The case of the National Institutes of Health (NIH) was presented by **Professor Scarpa**, Director of the NIH Center for Scientific Review (CSR). NIH annually funds grants totalling more than US\$20 billion supporting more than 200 000 researchers at over 3 000 research institutions in the US. It receives about 80 000 grant applications a year and engages about 18 000 reviewers. To adapt to changes in the research environment, changes were introduced: (1) communication with stakeholders was increased, (2) summary statements were made more uniform, and (3) electronic systems are being used to increase efficiency (electronic submission of applications and the use of knowledge management software to identify reviewers and assign application to them). More efforts are being undertaken to: (1) shorten the review cycle, (2) do more to recruit and retain more high quality reviewers, (3) decrease the burden on applicants and reviewers, and (4) improve the identification of significant, innovative and high-impact research. The use of electronic review modes is being tested (with the goal of having 10% of applications electronically reviewed by 2007). Other measures being considered to enhance efficiency include reducing the size of applications and shortening the review meetings.

Dr Bement Jr., Director of the National Science Foundation, presented the merit review used in selecting the proposals in his organisation. The NSF handles approximately 40 000 proposals a year (of which 10 000 are new grants) that are reviewed by approximately 50 000 volunteer referees.

The proposals are funded according to their merit, assessed against two criteria: (1) what is the intellectual



T. Scarpa

merit of the proposed activity? And (2) what are the broader impacts of the proposed activity?

The increasing complexity of science and engineering research and the broader mission of NSF (integration research and education etc.) increase the need for a diverse, inclusive and expanding pool of reviewers. A combination of researchers at an early stage in their career with more established ones, of individuals from different fields with a variety of intellectual perspectives, helps ensure that diverse viewpoints contribute to the process of identifying the best ideas.

The programme officers at the NSF and the advisory committees play an important role in ensuring high quality merit reviews. The programme managers (who either are permanent NSF staff or rotating from universities and other institutions), identify and manage merit review panels, recommend proposals for funding, balance a wide range of considerations in shaping the programme portfolio that they manage. The advisory committees (outside experts from industry, academia and other government agencies) to each NSF directorate and major office evaluate NSF performance and provide feedback on new directions and improvements to NSF programmes.

The ways to strengthen the merit review process at NSF were also discussed. They include: improving training for programme managers; building an expanding pool of diverse, highly qualified reviewers and developing new mechanisms to increase the transparency of the merit review process.

The review system at the Japan Society for the Promotion of Science was presented by **Professor Ono**, President of the Society.

In the fiscal year 2006 the JSPS received more than 88 000 applications, selecting about 22 000 of them. More than 4 100 document reviewers and 700 review panel members were involved during this process.

In 2003 the JSPS established a 'Research Center for Science Systems' which has three main functions: (1) to oversee application screening and project assessment for JSPS programmes, (2) to conduct surveys on science promotion policies and research trends, and (3) to provide the JSPS administration with recommendations on issues such as improving its screening and assessment functions and designing future funding schemes.

The centre is staffed with 113 programme directors and programme officers, top-level researchers in their fields, who work at centre part-time on three-year tenures. New officers are chosen from different institutions. Up to 2004, the reviewers were recommended by the Science Council of Japan and from 2005 onwards they have been chosen via an autonomous procedure at JSPS's Research Center for Science Systems. Reviewers are selected from a dedicated database that includes more than 40 000 potential reviewers. The JSPS has increased the number of reviewers so that they will review no more than 200 proposals each.

Mechanisms to ensure the quality of the selection process and clearly defined procedures to handle any conflict of Interest have been developed. To enhance transparency the screening policy and evaluation criteria are published in advance and reviewers' names are disclosed at the end of screening period. On request, unsuccessful applicants are provided with the following information: (1) their ranking within their subject category on a three-grade scale, and (2) the average score awarded by reviewers for each evaluation criterion.

Professor O'Reilly, Chief Executive of the Engineering and Physical Sciences Research Council (EPSRC) in the UK stated that peer review was introduced around 1690 as a means of vetting contributions to the Royal Society of London. Then he briefly presented the selection process at the EPSRC and explored the suggestion that peer review, as currently practised in research-funding bodies, may be unduly risk-averse.

The EPSRC handles about 5 000 application yearly and it involves a peer review college (whose members serve three-year terms) based on nominations by the research community and representative factors including gender and professional background. The peer review operations at the EPSRC are audited according to the quality management system ISO 9001.

Professor O'Reilly reflected on how to make peer review receptive to 'high risk/high (potential) return' proposals, which he defined as '*research activity where there is a considerable amount of uncertainty about the*

success, because completely new questions are being asked, new methods are used or "new" people are involved.' Examples were given of cases in which the current system may have failed to adequately address such proposals. The EPSRC established a separate 'adventurous interdisciplinary research fund' (1-2% of total available funds) devoted to proposals which may be characterised as '*high-risk/high uncertainty in the proposal but high impact if successful*'. To date over 600 proposals have been received.

Professor Hornbostel, Head of the Institute for Research Information and Quality Assurance, presented the results of an on-going survey of the members on the review boards of the Deutsche Forschungsgemeinschaft (DFG). Elected by the researcher community in Germany, the 577 members of 48 disciplinary review boards evaluate both the proposals of applicants and the reviews by referees from a scientific perspective (also serving to ensure the appropriate selection of reviewers).

The survey addressed a range of issues: from the perceived fairness of the peer review system to the perception of the burden imposed on the scientists involved. The partial results (based on responses from half of the initial sample) showed that the members of review boards are critical in judging the quality of reviews whereas, in most aspects investigated, the great disparity across scientific disciplines was striking. The respondents were in favour of rewarding the reviewers 'in their own currency (peer recognition etc.)', they regarded science as international and did not see any special advantages of involving more foreign reviewers. Against the background of the debate about the anonymity of reviewers, Professor Hornbostel concluded that establishing a review board seemed a good compromise between keeping the anonymity of the reviewers and the interest of the public. The respondents did not recommend the open review model (in which reviewers are known). The survey also showed that there was a strong call for open access to findings of funded projects by the respondents. Final results of the study are expected in April 2007.

2.3 Peer Review in Scientific Publishing

In scientific publishing peer review has been used for centuries to assess the merit of papers being submitted. Though a wide consensus holds that the system has proved its merit in assessing the validity of research outcomes, discussions on how it is (and should be) set up and managed are intense. In the plenary sessions, three speakers shared the experiences of their organisations, discussing arising pressures and the likely changes to come and how they are being addressed.

2. Plenary Sessions

Dr Campbell, Editor in Chief of *Nature*, discussed the pressures on peer review at *Nature* and other *Nature* journals. Unlike some other journals, *Nature* has no editorial board. The 26 editors who are chosen on scientific merit (strong publication track records, excellent labs) take direct responsibility for what is to be published, with the advice of (typically) two or three anonymous referees. About 60% of approximately 11 000 yearly submissions are rejected by the editors without review and within a week of receipt. The subsequent peer review process takes typically 4-5 weeks.

A survey among authors of articles in *Nature* published between June and November 2005 (with more than 1 500 respondents) showed that three quarters of them think that their paper was improved by the comments of reviewers. Dr Campbell discussed the 'healthy pressures' on peer review (including increasingly complex data sets, multidisciplinary and differing opinions of referees on the merit of a paper) and the 'unhealthy pressures' (including image manipulation, misconduct by author and referees, the use of reagents protected by commercial confidentiality) and how *Nature* handles them. He also discussed new approaches to an 'open model of peer review' and a trial made at *Nature*. In this model, articles not rejected at first sight by the editors were sent to confidential referees and in parallel posted on an open server for public comments. About 5% of authors, given the opportunity, opted for open peer review. An analysis of this model showed that most comments were not as useful as the solicited comments (less detail, less obvious expertise) and that the quality of most comments was below that of the solicited referees. *Nature* is currently running an online debate on peer review addressing current practices, ethical questions etc. (<http://www.nature.com/nature/peerreview/debate/index.html>)

Dr Simpson, Associate Editor of *Science Magazine*, described first the process of selection of articles at *Science Magazine*. After submission, the articles are evaluated by the board of reviewing editors whereby novelty and broader interest are the main criteria. In this process 70% of submitted articles are rejected and the rest is sent to reviewers. This second stage ends with rejection of about 70% of articles again and the rest being accepted (in some cases subject to revision). The overall, acceptance rate is between 6 and 8%.

The peer review, said Dr Simpson is a mechanism to ensure the rigour and accuracy of scientific research and as such serves also to give the public confidence in the quality of published results. This is even more important today when '*science is abutting ever more closely with core human values: stem cells, cloning and energy for the first time in many years of scientific history*'.

Measures to improve the quality of peer review include:

- Continuing to review and revise processes of peer review in journals
- Transparency: clarifying author contributions and potential conflict of interest
- Using measures to detect anomalies in data
- Exploring new models of peer review

Discussing the limits of peer review, Dr Simpson said that '*peer review assumes honesty and is based on trust. It can help establish scientific validity and rigour but it offers no guarantees.*'

Professor Mehlhorn, Vice-President of the Max Planck Society, observed that while modern Web technologies have changed the way scientific information is produced and distributed, the scientific community is starting to explore new options provided by the current technologies for organising the review process (which remains the key element for quality assurance in scientific publications). After a brief discussion of the weaknesses of current practices, he described new models of peer reviews which make use of new ICT advances and tools: the group review, the open review and the continued review.

Group review is used, for example, in computer sciences where refereed conference papers are the main means of information dissemination. Three to five months between submission and the conference the programme committee conduct Web-based discussions on the papers. Two to three members write an initial assessment that other committee members discuss further. This approach is seen as likely to identify breakthroughs and is seen as a fast process of quality assurance.

In the *Open review* process, the paper is available online for review by the scientific community at large. The advantages of this model are, among others, that the paper is likely to increase its quality as comment goes on, the high probability of recognising breakthrough and – because of the 'shame factor' – authors are not encouraged to post 'bad papers' and reviewers are likely to write good reports. The reviewers are rewarded by the fact that reviews are mini-publications in their own right.

In the *Continued Review*, the articles are continuously reviewed after publication. A track of citations and downloads is kept during this process, comments are continuously made by readers, and articles are amended by authors accordingly. The Living Review Series of the Max Planck Society (which includes Living Reviews in Gravitational Physics and Living Reviews in Solar Systems) are an example of how this model works.

3. Parallel Sessions

Peer Review: A Look from Different Angles

3.1. Assessment and Selection of Research Proposals: Challenges and Procedures in National Research Agencies [Sessions A1 and B1]

Peer review remains the central instrument used by research-funding agencies to identify which are the best research ideas to support. However, given the changes in research practices, together with the increasing number of research proposals received and organisational changes in national science systems, the practical approaches to peer review must also evolve to reflect these developments.

The conference reserved two sessions for discussing how national research agencies organise their peer review in practice and address the challenges they face. The sessions (A1 and B1) were chaired by Professor Václav Pačes, President of the Academy of Sciences of the Czech Republic and Dr John Marks, Director for Science and Strategy at the European Science Foundation, respectively.

Session A1

Chair: **Václav Pačes**, *Academy of Sciences of the Czech Republic*

Henrik Bruun, *University of Turku*, Evaluating interdisciplinary research proposals: the case of the Academy of Finland

Oh-Kab Kwon and Jong Geon Lim, *Korean Science and Engineering Foundation*, Reorganisation of the evaluation system at KOSEF, Korea

Pavel Chráska and Josef Syka, *Czech Science Foundation*, Peer Review in the Czech Science Foundation – experiences and challenges

Nüket Yetis, Omer Cebeci and A. Yavuz Oruç, *Scientific and Technological Council of Turkey*, Effective evaluation and funding of research projects

Rapporteur: **Bernard Avril** (ESF)

Speaking in the session A1, **Professor Bruun** (University of Turku, Finland) noted that interdisciplinary research had received renewed attention in recent years but that few funding schemes have been specifically developed to support interdisciplinary research.

This has led to some concern regarding the appropriateness of the current peer review procedure for its evaluation. One of the main arguments presented was that peers use disciplinary criteria for making their assessments, thus creating a potential bias against interdisciplinary research. Professor Bruun presented part of a study commissioned by the Academy of

Finland¹ to investigate to what extent it promotes interdisciplinary research. The study covered the years 1997, 2000 and 2004 and analysed the General Research Grant Scheme, a responsive-mode funding scheme of the Academy.

Looking at the type of research the Academy of Finland funds, the study found that 42% of the projects were either multi- or interdisciplinary. The study also showed that the interdisciplinary nature of research proposals did not seem to influence their success rate, which was approximately 20% for both single disciplinary and interdisciplinary research proposals. The authors concluded that the evaluation system used by the Academy of Finland was efficiently handling interdisciplinary proposals.

Discussing the results, Professor Bruun commented that they were seen to be consistent with the outcomes of international research on the funding of interdisciplinary science. He also added that it was important to distinguish between ‘interdisciplinary research’ and ‘radical science’ (i.e. unconventional, risky or adventurous science). The findings challenged the commonly held view that reviewers can be biased against the innovation of interdisciplinary proposals. It seems that that interdisciplinarity is a more common phenomenon in science than is generally believed, suggesting that it has not been marginalised or even slowed down by the peer review system.

The process of the reorganisation of the Proposals and Programme Evaluation System currently undertaken by the Korean Science and Engineering Foundation (KOSEF) was presented by **Dr Kwon**, its Chairman and Chief Executive. In its efforts to improve the country’s standing in science and technology, the research system in Korea has undergone a massive restructuring, in the course of which KOSEF saw its activities expanded (from basic science research to large-scale R&D programmes) and its annual budget increased (to approximately US\$1.5 billion in 2006).

In this context, KOSEF needed to reorganise its evaluation systems to accommodate both basic research programmes and national R&D programmes. A Standardised Evaluation Process (SEP) was introduced, covering the pre-funding, progress and post-funding evaluation. In addition, a new system for categorising research programmes was developed, based on budget size, the project duration and competition rate, with four classes, Horizon 1 to Horizon 4. Research programmes categorised as Horizon 1 have substantial budgets (about US\$ 1 million per year), long time frames (about nine years) and low selection rates (only about 15% of

1. Bruun, H., J. Hukkinen, K. Huutoniemi, J. Thompson Klein (2005). *Promoting Interdisciplinary Research: The Case of the Academy of Finland*. Academy of Finland Publication Series 8/05. Helsinki: Academy of Finland.

3. Parallel Sessions

proposals funded), while Horizon 4 programmes have low budgets (US\$15000 a year), short time frames (about two years) and high success rates (about 75% of funded proposals). Today on average 39 referees are involved in evaluating Horizon 1 programmes and 12 in Horizon 4 programmes. Although KOSEF aims to reduce the number of referees to 20 and seven respectively, one of the main challenges it faces is the selection of good referees.

KOSEF is considering moving from the 'open model' of referee management to the 'closed model' in which referees are contracted to work for a defined period of time. The Japan Society for the Promotion of Science (JSPS) and the UK Engineering and Physical Sciences Research Council (EPSRC) were cited as examples of organisations using such a model. In those organisations a pool of official referees is contracted for two and three years respectively (containing about 4800 and 4000 referees). A system of 'mileage', to provide an incentive for reviewers, is currently under discussion. For example, the 'miles' could be used to support the referees' participation in foreign conferences. KOSEF will also appoint a research planning and evaluation committee consisting of outside members to monitor the reviewers and advise KOSEF on how to continuously improve its review process.

Finding and managing high quality referees is also a problem faced also by the Czech Science Foundation (GAČR) as was illustrated in the presentation of **Professor Chráska**.

The Czech Science Foundation (GAČR) was founded in 1993 to fund basic research in all disciplines. For the year 2005 the budget was €48.9 million and 2843 grants were awarded, of which 2145 were from the Standard Grant Scheme and the rest were grants for individual postdoctoral scholars and postdoctoral teams.

It has been observed that with small national scientific communities, conflicts of interest are difficult to avoid and the related concerns of patronage and clientelism are more frequent.

The GAČR is addressing this problem by annually changing one third of the members of scientific committees and subcommittees (those committees oversee the review process and make recommendations on funding to the board of the foundation) and by having proposals submitted by members of those committees evaluated and ranked by a special committee. Additionally, the GAČR has been using international referees from the beginning and requests that proposals be written in English. Yet it is *'getting more and more difficult to find high quality responding reviewers'* said Professor Chráska. About 24% of the 3585 contacted national referees in the year 2005 did not submit their reviews. The proportion among foreign referees was even higher: with 43% (of the 3673 approached) declin-

ing to take part in the review process. One solution discussed at the GAČR would be to offer a financial incentive to the referees. Yet paying €50 per submitted review would increase the administrative costs of the foundation – which is currently at 2.6% of the total budget – by about 50%.

Another issue discussed was the comparability of reviews undertaken by foreign and domestic referees. An analysis of the evaluation indicated that the Czech reviewers are more likely to deliver a 'softer' evaluation than their foreign colleagues.

Professor Oruç informed the conference participants that the Scientific and Technological Council of Turkey (TÜBİTAK) is considering involving scientists from outside Turkey in its review panels to overcome what is seen as the major shortcoming of the panel system: a limited pool of reviewers and panellists. This system, used by TÜBİTAK since 2004, is modelled on the panel review system of the US National Science Foundation (NSF). Four notable differences were highlighted:

- TÜBİTAK panel system relies more heavily on panel recommendations than the NSF panel system.
- TÜBİTAK panel system uses a more quantitative (numerical) evaluation based on a 3-dimensional evaluation scale: (a) intellectual merit, (b) broader impact, and (c) feasibility.
- TÜBİTAK panels are run by non-resident scientists and researchers unlike NSF panels which are moderated by permanent programme officers or scientists coming from academia and temporarily appointed to the NSF (also called IPAs because their appointment is made under the Intergovernmental Personal Act).

Funding decisions are more centralised and made at higher levels of administration, unlike NSF funding decisions that are carried out almost always by programme officers.

The panel system currently operated by TÜBİTAK is favoured by most scientists in Turkey. A survey conducted in 2006 (results of which were being analysed at the time of the presentation) shows that irrespective of funding, publication records and geographical locations, researchers in Turkish universities are convinced of the efficiency of the current system (1441 researchers participated in the survey). Increasing the participation of international panellists was considered to be the key means of improving the current system, not least because of the increasing number of proposals evaluated. Between 2004 and 2005 both the number of proposals evaluated and the average number of the proposals per panel doubled from 1447 to 3401 and from eight to 16 respectively. In his concluding remarks, Professor Oruç proposed to carefully review the issue of the optimal panel size and associated cost.

Session B1

Chair: **John Marks**, *ESF*

Christoph Schneider, *Deutsche Forschungsgemeinschaft*, The worst form of deciding on grant proposals except all others - an administrator's view on peer review

Minghong He, *National Natural Science Foundation China*, Discussion on peer review of grant applications in NSFC

Dong-Pil Min and Dongseob Kang, *Korean Research Foundation*, Assessment system for grant proposals in Korea

Rapporteur: **Alexis-Michel Mugabushaka** (ESF)

In Session B1, three speakers shared their experiences in the practices and challenges of peer review in their respective organisations: the Deutsche Forschungsgemeinschaft in Germany (DFG); the National Natural Science Foundation of China (NSFC) and the Korean Research Foundation (KRF).

Starting the session, **Dr Schneider** from the Deutsche Forschungsgemeinschaft (DFG), identified two major problems the peer review system is inherently confronted with: *'On one hand as it is a part of a decision-making process that has to produce clear results, on the other hand it operates under conditions of high uncertainty'*.

As the peer review process is central in allocating rather scarce resources (research funds, journal space, recognition), it stirs strong feelings. The main criticism the peer review system faces is that, *'being a social activity'*, it involves human beings who may fail to recognise the true merit or simply have no time to properly review the grant. In a 'thought experiment', Dr Schneider identified six mechanisms to allocate research funding which could be alternatives to peer review systems as currently practiced:

- (1) Decision by bureaucrats
- (2) Decision by political correctness
- (3) First come, first served strategy
- (4) Lottery
- (5) Decision by indicators
- (6) Enlightened absolutism.

Reflecting on the problems they cause, he concluded: *'my conviction, like many other practitioners in research administration, is that there is no alternative to peer review'*. The challenge is not to replace the existing system but to improve it and to overcome its weaknesses and drawbacks. Key issues are:

- a proper management of the process
- establishing sound safeguard mechanisms to minimise bias and conflict of interest.

The process to award research grants at the National Natural Science Foundation of China (NSFC) was described by **Dr He**. The NSFC is the major research funding agency in China. In recent years, it saw an annual increase of proposals of 20% (to reach 64 730 proposals in 2006). The NSFC has a dozen research funding schemes with slightly different selection procedures and criteria. Yet the basic system of proposal selection typically consists of three stages. After the submission, the proposals are checked by NSFC staff and between 95 and 97% of the proposals enter the second stage of 'mail review'. This is the most critical step involving contacting two to five experts (with 95-98% feedback from the reviewers). The programme managers critically choose the reviewers from a 60 000 record database. The reviewers have not only to be experienced experts in the given research field and be still working actively in the field and familiar with the frontier of the field in question but also 'tolerant' of different research approaches. About 30-40% of the proposals pass this step to be reviewed by one of the 60 panels of the NSFC (involving nine to 14 panellists). The review panel makes the final decision by majority vote (mostly anonymously). The success rate lies between 15 and 25%. Each step in this selection process has its own safeguard mechanism to avoid bias. The panellists and programme managers at the NSFC are encouraged to spot 'risky project proposals' which are then funded under stricter conditions: their duration is shorter (e.g. one year) and their budget is smaller (roughly a third or a half of the normal project budget). About 3% of the total budget is devoted to such projects.

The next presentation was made by **Dr Kang**, Director of the Innovation and Strategy Division at the Korean Research Foundation (KRF). The KRF is the major research funding agency for universities in Korea. In 2005 it had a budget of US\$1 billion and handled 16 181 proposals (of which about a quarter were successful). From 2006 onwards, the KRF will be using a new system ('the expert driven review') to assess research proposals. Traditionally, KRF has been selecting the research grants using 'panel reviews' only. Researchers submitted their proposals through an online system (One Stop Research Management System) similar to the NSF Fast Lane. Programme managers at KRF established panels of six to nine reviewers who elect their chairperson. Each proposal was assigned to two panellists (a speaker and a discussant) who presented and discussed their assessment of the proposal to the panel. The chair and his panel gave marks to the proposals and sent a completed review form to the KRF. The funding decision was made by the Grant Selection Committee which is made up of KRF senior and mid-rank officials and officials from the Ministry of Education and Human Resources Development. The committee published a

3. Parallel Sessions

preliminary list for public comments on its homepage to prevent duplicated or overlapping supports. After a week without public complaints the funding decision became effective.

In contrast to this 'panel only system', the new system of an 'expert-driven review' – said to be similar to the 'mail+panel' system of the NSF in the US – will involve an assessment by three reviewers who are to be recognised scholars in the research field of the proposal. In the second step the pre-selected proposals will be reviewed by a panel organised by the programme manager.

Discussions [Session A1 and B1]

In the discussions from the floor it was observed that although the speakers presented practices of organisations situated in different parts of the globe and from countries with different research-funding systems, the practical mechanisms in their peer review system are essentially similar. To **properly manage** the process of peer review all the organisations presented have mechanisms to rigorously check potential sources of bias and of conflict of interest. In the organisations presented, a multistage review process involving review boards was introduced.

Discussing models to ensure that the present system does not disadvantage the **'risky research'**, the model used by the NSFC was seen as noteworthy. The panellists and programme managers at the NSFC identify 'risky project proposals' and recommend them for funding under stricter conditions.

Another approach discussed by participants was 'to take the risky research out of the peer review mechanism'. For example in identifying teams and labs that are already successful and allocating them extra funds for adventurous research.

In the discussions it was pointed out that the present system should also be credited for catching a great many 'risky projects'. The research and science is *per se* about new knowledge and is adventurous by its very nature. From this perspective the current peer review system helps to advance it and might be less biased towards innovation than generally perceived.

The presentations about the practices at the National Natural Science Foundation of China and the Korean Research Foundation touched upon the **IT tools that support the peer review systems**. To cope with the increasing volume of proposals, those organisations have developed tools for the reviewers to access the proposals and rate them online. And the selection of the referees is supported by large in-house databases containing 60000 records at the NSFC and 31 000 at the KRF. Additionally the KRF uses the Academic Researcher Database, a system containing data

on 160000 researchers including university professors in universities and institutes. In the discussions it was pointed out that, though the computer systems are important and helpful, the work done by the programme managers in identifying and selecting the right reviewer is most critical. There was an agreement that *'people are more important than computers'* in this process.

'Finding the best reviewers' will remain the greatest challenge the research managers face. The discussions echoed the view that indeed reviewers must be outstanding researchers in their field. Yet this criterion alone may not be enough to meet the requirements of high quality peer review. It is also equally important to choose reviewers who are open to other ideas and tolerant of scientific approaches they may not necessarily share.

3.2. Assessment and Selection of Research Proposals in International Organisations and Transnational Funding Schemes [Sessions A2 and B2]

Behind the concept of the European Research Area lies the vision of overcoming the fragmentation of Europe's research efforts. The vision is being realised by research collaboration in pan-European research organisations and in numerous initiatives by national research organisations seeking to catalyse research collaborations and network among the main stakeholders. They all face a challenging problem of finding a system to assess the quality and potential of research proposals coming from researchers used to different assessment rules and procedures. Two sessions of the conference looked into practices and problems in evaluating research proposals in the international context. Session A2 was chaired by Professor Omling, Director General of the Swedish Research Council.

Session A2

Chair: **Pär Omling**, *Swedish Research Council*

Les Grivell, *European Molecular Biology Organization*, Peer review at EMBO

Francesco Fedi, *European Cooperation in the Field of Scientific and Technical Research*, Assessment, monitoring and final evaluation of COST Actions

Liz Allen, *Wellcome Trust*, Looking for landmarks – the importance of post-award peer review

Rapporteur: **Arja Kallio** (ESF)

The peer review system as used by the European Molecular Biology Organization (EMBO) was presented by **Professor Grivell**. Established in 1964, the mission

of EMBO is to promote excellence in molecular life sciences in Europe. It is a membership-based organisation and members are elected annually on the basis of their record in research. In 2006, EMBO had 1 100 EMBO-elected members in Europe (including 38 Nobel Prize winners) and over 60 associate members globally. EMBO uses peer review in its two main activities: research funding and science publications. For both, to a large extent, EMBO uses the expertise of its members.

EMBO funds, among other things, postdoctoral and long-term fellowships for outstanding scientists in life sciences. In 2005 about 200 awardees were selected from 1 235 applicants. Applications for long-term fellowships are pre-screened by the members of the programme committee (with a balanced scientific, geographic and gender composition, where possible) and further evaluations and individual interviews are carried out by EMBO members and Young Investigators. Final decisions are made after discussions and individual scoring of applications.

A particular issue to which EMBO has paid attention is the potential bias of peer review against female applicants in its funding schemes. A pilot evaluation scheme with gender-blinding was undertaken.

Professor Fedi, President the Committee of Senior Officials of the European Cooperation in the Field of Scientific and Technical Cooperation (COST) presented the process of quality control used in COST.

Founded in 1971, COST is Europe's first intergovernmental European network for the coordination of nationally funded research. It has 35 member countries. It supports the cooperation efforts of European research through COST Actions. The three pillars of quality control at COST are: assessment of proposals for new Actions, monitoring of the running Actions and final evaluation of completed Actions.

Proposals for new Actions are submitted in a two-stage process: preliminary proposals can be submitted at any time and on any subject (two collection dates per year). After a formal check the proposals are pre-screened and ranked according to specific criteria by COST Domain Committees involving at least eight members. Subject to the available budget for new Actions, the highest-ranked preliminary proposals are invited to submit a full proposal. Full proposals are evaluated by an External Expert Panel, which is coordinated by rapporteurs designated by the relevant COST Domain Committee (DC). Each DC establishes a ranked list of proposals and a meeting between DC chairs establishes a general ranking list with recommendations on funding. The CSO (Committee of Senior Officials), which is the highest decision-making body made up of representatives of all COST member countries, takes the final funding decisions.

In response to a call issued in May 2006 for the selection of 25 new Actions, 824 preliminary proposals were submitted.

Dr Allen's presentation focused on the methods used by the Wellcome Trust to evaluate the outputs and achievements of the funded projects. With a budget of about £450 million, the Wellcome Trust funds about 6 000 researchers in over 50 countries. The selection of the projects is done on a peer review basis. In 2004/05, about 2 500 new applications were submitted and 800 awards granted after a peer review process in which over 10 000 referees were contacted. With so much effort being put into getting the proposals properly selected, the post-award evaluation at the Wellcome Trust focuses on the questions whether the right choices were made and whether some funding mechanisms are more likely to yield more important research results than others. The Wellcome Trust has used the bibliographic system Pubmed which indexes papers acknowledging the Wellcome Trust since May 2005. About 1 000 papers recorded for the period between May 2005 and September 2005, were reviewed by a 'peer review college' to assess their importance to the research community.

Categorising the papers into four categories ('landmark research', 'major addition to knowledge', 'useful step forward' and 'for the record'), the peer-review college found that around 9% of the papers can be considered at least as a 'major addition to knowledge'. It was also shown that particular funding schemes are related to 'success'; for example larger grants seem to be more prone to yielding 'more important' research. Dr Allen concluded by calling for cooperation between research-funding agencies to conduct cross-agency evaluations of this kind for benchmark purposes and to work jointly towards commonly acknowledged practices and conventions.

Discussions [Session A2]

Discussing the approaches taken in the organisations which shared their experiences, the participants were in agreement that the quality of the peer review is determined by the quality of the reviewers and that the acceptance of the outcome it produces depends heavily on the level of the trust that those who are reviewed have in their 'peers'. Therefore a thoughtful selection of the 'right' reviewers is the critical aspect of any peer review exercise.

Research networks in their setting and *modus operandi* differ from the 'classical' research project carried out at single institutions. The international dimension in the projects presented adds even more to their complexity. In the discussions it was suggested that the reviewers who are to evaluate the international research networks be 'educated' in adequately assessing their

3. Parallel Sessions

quality and potential. Clear guidelines and criteria were seen as crucial. They should be developed and given to the reviewers before the evaluation task.

Other challenges mentioned include gender balance and ethical considerations. It is an increasing problem that women scientists are overloaded with requests to serve as reviewers. Gender balance in review panels was, however, considered to be of the utmost importance and something to be definitely aimed for. Ethical guidelines, for example concerning conflict of interest, should be clear and followed during peer reviews.

Panel Discussion on Transnational Research Funding Programmes [Session B2]

Chaired by Professor Andersson, Chief Executive of the European Science Foundation, this session discussed the practices and problems in peer review in various transnational research funding initiatives.

Session B2

Chair: **Bertil Andersson**, *ESF*

Raimo Väyrynen, *Academy of Finland*

Izo Abram, *Centre National de Recherche Scientifique*

Rudolf Novak, *Austrian Research Fund*

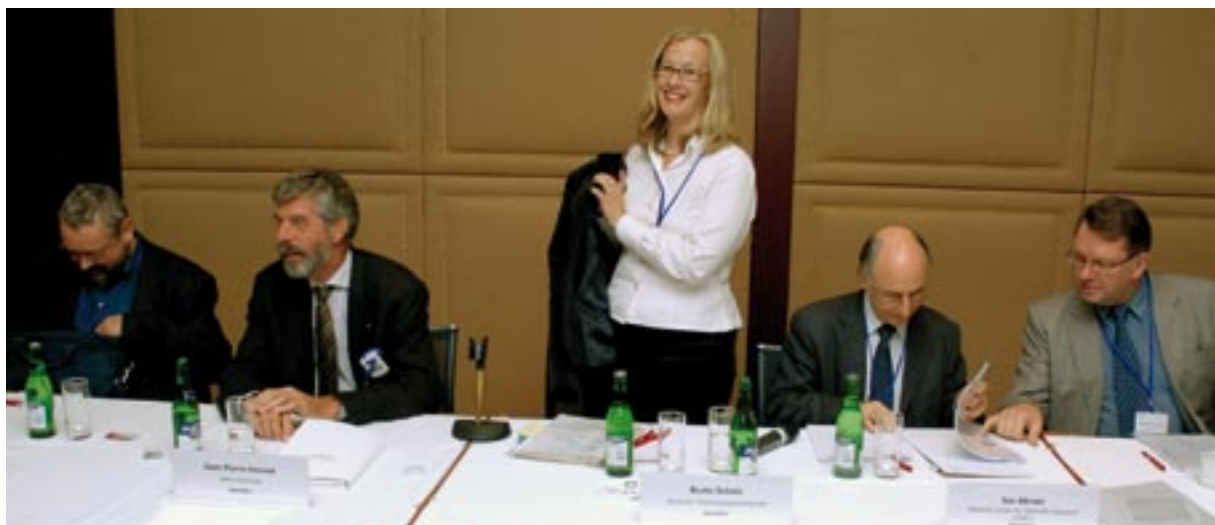
Jean Pierre Henriët, *Ghent University*

Beate Scholz, *Deutsche Forschungsgemeinschaft*

Rapporteur: **Daphne den Hollander** (NWO)

Taking a look at the peer review systems in European Nordic Countries, **Professor Väyrynen**, President of the Academy of Finland, noted that the science systems in the various Nordic countries are different indeed, but this does not prevent the countries from working together. Under the framework of the Nordic Cooperation developed since the 1960s, national science organisations can choose activities of interest to address jointly with their counterparts in other Nordic countries. In some areas, joint research has been funded, while in other areas they worked together more on a policy level. *'It's all about identifying common concerns and being flexible in what you want to achieve together'*. When it comes to peer review, there are experiments in working closely together and in using each other's experts. The countries are small, and there is a lot of pressure on the peer review system. Everyone is 'fishing from the same pool'. There are concerns that with advent of the European Research Council (ERC), the competition for best experts to act as referees will intensify.

Dr Abram, Director of the international department of the French National Centre for Scientific Research (CNRS) presented case studies of peer review in international research activities. The ERANET NanoSciERA is a cooperation between 12 countries who have recently launched a joint call for proposals with a 'common pot' mechanism to fund the best projects. To combine different national practices and get to a procedure that allows everyone to participate is a very delicate and important task not only for the efficacy and efficiency of the process but also for its acceptability of the transnational



From the left: R. Novak, J.-P. Henriët, B. Scholz, I. Abram, R. Väyrynen

process. Researchers, knowing their own home institution, project onto it the trust they have in their home agency. In developing the peer review of the NanoSciERA, the organisations involved faced two problems: to develop principles to prevent conflict of interest and to assess the level of expertise of the referees approached.

Regarding the conflict of interest (or perceived conflict of interest), the NanoSciERA adopted the guidelines of the Deutsche Forschungsgemeinschaft (DFG) because they had a very useful written document explicitly relating to the appearance of conflict of interest. Furthermore, panel members from outside the consortium countries were selected (on recommendation of the participating organisations). But it is still difficult as *'the best people know the best people (also from different regions of the world)'*. In order to further ensure a fair and competent assessment, the referees were asked to self-rate their expertise.

Further features of the review system include a software-tool developed for referee assignment on the basis of key words and giving the applicants the opportunity to reply to referees' comments ('rebuttal'). Dr Abram concluded by saying that notwithstanding the fact that the ERA-NET scheme may be too new to say if it will establish itself as a means of transnational collaboration, a sound peer review system is a key factor for its acceptance. In ERA-NET, evaluation and funding are closely linked. The outcome of the peer review is the distribution of funds.

The European Young Investigator (EURYI) Awards was created in 2003 by the EuroHORCs in collaboration with the ESF.

Dr Scholz, in her capacity as Chair of the EURYI Management Committee presented the most important features of the EURYI scheme. Currently there are 15 countries participating in the scheme. The selection of the awardees (who respond to a call issued by the ESF) is organised in two steps. The first step (S1) in selection is in the hands of national research councils and the second step (S2) is done by an international panel convened by ESF.

A major challenge in establishing this selection mechanism was to ensure that the funding agencies involved follow similar processes in the first phase (S1). As different organisations have different approaches: *'It is essential to look at the commonalities and to get as close as you can'*.

The participating organisations agreed on the criteria to be used in the first stage: (1) for the candidate: potential and achievements, (2) for the proposal: scientific quality and originality, (3) for the host institution: level of excellence and commitment.

They also agreed on the main lines of procedures to follow: (1) using three expert referees, (2) sifting the

proposals by disciplinary panels, (3) a shortlist to be submitted to the international panel.

Professor Henriet, who is member of a EUROCORES Review Panel, sees two pillars of wisdom in Europe: on the one hand the European Commission funding top level science with political as well societal relevance, on the other hand ESF or 'research council cooperation' funding top level science, curiosity-driven and entrepreneurial in its character.

A strong point of the EUROCORES scheme is a generally high level of commitment and flexibility (shown by funding agencies) to make the best out of it. A weak point, however, is that expectations are systematically not met in terms of funding and in terms of rules/regulations if one or two funding bodies drop out or cannot fund all projects. This is very difficult to understand, and good communication between funding bodies and the ESF, as well as between funding bodies, ESF and scientific communities is essential.

Taking a comparative perspective on the ERA-NET and EUROCORES schemes, **Dr Novak**, Head of Strategy Department at the Austrian Science Fund, emphasised the great similarities between them. The ERA-NET might aim for a common pot, where the peer review process and funding of projects is linked. In EUROCORES, these two processes are separated. Separation causes confusion for scientists.

Three issues are very important for acceptance of peer review results: (1) standardised procedures, (2) common cultures, (3) transparency and liability.

In the light of this analysis Dr Novak sees room for improvement of the EUROCORES scheme. The procedures are rather clear but complicated (multi-step); it is not an easy task to coordinate national funding agencies and at the same time to develop quality standards for all participants across borders, a challenge that people outside Europe may find difficult to understand.

Attention needs to be paid when the decision of the international review panel is not understood by the national funding organisation. It is recommended that communication be clear about financial frames so that scientists can calculate the potential for receiving funds. Communication is essential in the theme selection, in the review process, in communication to scientists about funding.

Discussions and Recommendations [Session B2]

The key issues are to define a set of best practices based on different national peer review systems and to develop trust in international peer review. The initiatives presented in the session (ERA-NET, EUROCORES and EURYI) all had to meet the challenge of combining different national practices and developing procedures which would be accepted by participating organi-

3. Parallel Sessions

sations. Their experiences show that transnational cooperation, though complicated, is not impossible. Their experiences show the way to:

- Combining different sets of procedures (use existing practices) and being flexible
- Developing a set of minimum of standards agreed upon between organisations (the critical issue has been to 'find a common ground without losing the quality standards')
- Further developing 'national' procedures when engaged in transnational programmes.

The following issues were identified as crucial in the future of transnational cooperation in research funding:

- Organisations are fishing in the same pool of experts; with the advent of the European Research Council, this may become worse
- Trust and reliability: the scientific community needs to accept the peer review process. For this, solid procedures are needed and only the best experts should be engaged in the peer review process
- Taking transnational cooperation further by pooling evaluation systems across Europe together.

The session concluded with recommendations for the EuroHORCs and ESF:

- To work towards a common pool of referees, so that organisations do not all fish from the same pool, but organise 'fishing trips' together
- To take it a bit further: link knowledge systems, so that an immense new source for scientists is created
- To develop guidelines for a minimum set of standards (which can be agreed upon) to 'find one another' in a European procedure and ensure trust and liability. Best practices from different organisations would help
- In the future, mechanisms such as ERA-NET and EUROCORES should be combined. EuroHORCs, as active ERA-NET and EUROCORES participants, should elaborate on this and make a statement. It was remarked that (separate) evaluations of ERA-NET and EUROCORES are now taking place: a more strategic evaluation at a more conceptual level, not of a specific 'mechanism', but of the underlying principles and ideas would be useful.

3.3. Peer Review in the Evaluation of Research Institutions [Sessions A3 and B3]

Following a general trend in the public sector, in the last decades publicly funded research institutions have witnessed a growing demand for accountability. They meet this demand by periodic assessment of their performance, carried out in most cases by external review panels. Two sessions discussed the principles and the practices of the peer review system as used

today in the evaluation of research institutions. The sessions, chaired by Professor Makara, President of the Hungarian Research Foundation OTKA (A3) and Dr Gudmundsson, Director of the Research Council of Iceland (RANNIS) (B3) respectively, addressed the methodological approaches and presented case studies of the use of peer review in the evaluation of institutions.

Session A3

Chair: **Gábor Makara**, OTKA

Reinder van Duinen, former ESF President, Review of research institutes and programmes; experience and reflections

Jean Luc Devenon and Jean Yves Perrot, French Research Institute for Exploitation of the Sea, Evaluation of research units at IFREMER

Antonio Bertin, Istituto Nazionale di Fisica Nucleare, The evaluation experience of the Italian National Institute for Nuclear Physics: state-of-the-art and open problems

Thibaut Lery, Irish Centre for High-End Computing, Avoiding double peer review in the provision of scarce infrastructure

Rapporteur: **Monique van Donzel** (ESF)

Dr van Duinen, former President of ESF and NWO in the Netherlands, reflected on his experience with the evaluation of research institutes in the Netherlands. He noted that it is important to realise that in this context more is at stake than just scientific excellence. Is 'peer review' the right phrase to use at all?

Publicly funded research institutions often have multiple tasks: they perform high quality scientific research and provide infrastructure and services to other institutions. In their evaluation the review panels are expected to assess not only the quality of the research outputs but also the quality of the service provided and the management of the institutes (at the level of the whole organisation and as well as at the research unit level). The central questions for the review of institutes are: are we doing the right things and are we doing the things right? The objectives of such a review are aimed at improving the quality of the research, the research management and leadership as well as the accountability. Another dimension, equally important, is the fact that the assessments of research institutions cover past performance and future plans. Confidence about the future is based on past performance.

Because such assessment exercises are often seen as a burden on the management and the staff, mechanisms have to be in place to organise the assessment itself in way which reduces the burden and ensures a smooth evaluation process and avoiding fossilisation.

Discussing the possible outcomes of an evaluation, it was suggested that organisations should avoid directly coupling the results of the assessment and the funding decisions. The outcome will have an effect on the institution but this should be done in an indirect way. A buffer between the assessment and the funding is needed.

Professor Devenon presented a concrete example by explaining how the evaluation of the research units at IFREMER was organised. IFREMER (French Research Institute for Exploitation of the Sea) aims at ensuring better knowledge, assessment, value enhancement and streamlining in the exploitation of marine resources and to improve knowledge and means to protect and restore the marine environment. In addition it creates and manages facilities of national interest (fleet) and it gathers, disseminates and enhances national and international oceanographic information.

The evaluation of the research units and laboratories at IFREMER are conducted by an external expert group, on the basis of the analysis of a report of assessment and prospective written by the research unit to be evaluated and also on the basis of the conclusions of a visiting committee composed of some of the members of the expert group.

The dual mandate of dealing both with excellent research and with providing a facility and services – as clearly shown in the mission of IFREMER – proved to be a real issue in the assessment. There are currently no performance indicators for such services. This was perceived as a problem. The separation of the assessment of the scientific excellence and of the service provided is crucial.

Professor Bertin gave another example of an assessment, and focused on the scientific excellence of the Italian National Institute for Nuclear Physics (INFN). The INFN is an organisation dedicated to the study of the fundamental constituents of matter and conducts theoretical and experimental research in the fields of sub-nuclear, nuclear and astroparticle physics. Fundamental research in these areas requires the use of cutting-edge technologies and instrumentation, which the INFN develops both in its own laboratories and in collaboration with industry.

In 2005-2006 a complete evaluation exercise of the scientific productivity was done for the first time, including a set of quantitative indicators relevant to measuring the impact of publications. Also introduced was a year by year monitoring of national and international comparison with other disciplines.

The outcome of the INFN evaluation was positive. However, it was accepted that improvement can be achieved through measurable results; the inclusion of numerical measuring needs to be considered for the future.

Dr Lery presented the case of evaluating institutions that give access to large facilities. The real difficulty is to define the correct criteria to evaluate the service provision that can deal with research, services and teaching. The criteria have to take into account quality, productivity, accomplishments and also relevance and management. Of particular importance in this case was the issue of timelines. With a typical timeframe of two to three years between evaluations for institutions, a yearly evaluation for national facilities and quarterly plus annual reviews for international facilities, revisions and changes are not necessarily made in time to be included in following reviews. Examples at Irish and European levels have been given to show how institutions overcome such a difficulty. The issues of multiple stage evaluation, comparability and reduction of the amount of paperwork were raised. All these issues need to be addressed in order to make evaluations more effective.

Session B3

Chair: **Hans Gudmundsson**, *RANNIS* Island

Pär Omling, *Swedish Research Council*, Linnaeus Grants – university funding in national competitions

Nicole Haeffner-Cavaillon, *INSERM*, Peer review and bibliometrics

Rafael Rodrigo, *Spanish Council for Scientific Research*, The role of institutional evaluation in the CSIC strategic planning

Rapporteur: **Nina Kancewicz-Hoffmann** (ESF)

In Session B3, all three presentations concerned cases of evaluations of institutions, their contexts, procedures and results. The first presentation described the first competition for Linnaeus Grants where universities applied for funding for selected research fields. The second presentation described the use of bibliometric indicators, along with peer judgment, in evaluating research teams within INSERM, the French health research institute. The third presentation reported on the review of CSIC, the largest public research-performing institution in Spain, which is undergoing a reorganisation. The aim of the review was to develop a strategy for the new, reorganised CSIC.

Professor Omling, Director General of the Swedish Research Council presented the peer review procedures developed for a new funding scheme: Linnaeus Grants. Introduced in 2005 and based on a principle of 'Support to Strong Research Environments', the Linnaeus Grant scheme, funded from a share of block grants (institutional funding directly disbursed to the universities by the government), funds university research based on national competition between the universities. The

3. Parallel Sessions

Swedish Research Council had to develop a new peer review system adequate for this scheme.

In this scheme, universities (vice-chancellors) can apply for long-term substantial financing (up to 10 years and up to €1 million per year) in a specific research area. Large universities can submit several applications. The universities are required to submit applications in agreement with their strategic plans, to lay down a communication strategy and policies regarding ethical issues, gender balance etc. They also have to commit to co-funding the initiatives (at least 50% co-funding is expected) but they may include already existing resources in the co-funding.

The assessment criteria focus on scientific quality attained and the potential for scientific renewal. The level of commitment of the applicant university is also taken into consideration.

The selection is made by international panels: four subject-oriented panels and one general expert panel, which also takes the final decisions. Only foreign reviewers participate in the evaluations and decisions to avoid conflicts of interest as all Swedish universities are participating in the competition.

The procedures used in evaluation of laboratories were presented by **Dr Haeffner-Cavaillon**, Head of the Bibliometric Unit, at INSERM. Created in 1964, The French National Institute of Health and Medical Research (INSERM) aims at developing closer links between 'basic' and 'clinical research'. It has a budget of about €600 million (in the year 2005) and has over 6 000 scientists working in biological, medical and public health research. The evaluation is conducted by scientific committees (Commissions Scientifiques Spécialisées, CSS) with participation of external, also international, experts. The evaluation involves three stages: site visits, ranking and interview of the unit leader.

Bibliometric indicators are used as a support tool for the evaluation committees. Several factors are used because each indicator taken individually would cover only one facet of the performance. At INSERM, the following indicators, among others, are considered: number of papers (ISI database of the Institute for Scientific Information); number of citations; number of Top 1% and Top 10% papers, number of papers per researcher, journal impact factor. The presentation discussed in detail each indicator and its possible use in the evaluation process.

The presentation concluded by remarking that the bibliometric indicators should be used only as complementary to peer review judgment and cannot replace it. The experience at INSERM shows that introducing bibliometric indicators does not dramatically change the results of an evaluation. However, it strengthens the transparency of decisions and provides for their

substantiation. It was at the same time noted that the bibliometric indicators should be used with an awareness of their advantages and limitations

Professor Rodrigo presented the review process at the Spanish Council for Scientific Research (CSIC) of which he is Vice-President. CSIC is the largest public institution in Spain devoted to multidisciplinary scientific and technological research in all research domains. It has a budget of about €600 million (in the year 2005) and employs about 2 500 research staff and 4 500 postgraduate students and postdoctoral scholars. It is organised into 116 institutes and 134 associated units collaborating with universities and other research institutions.

For the development of the CSIC multi-annual strategic plan for the period 2006-09, institutes were asked to draw up their own Strategic Plan Proposals which were assessed by external peer review panels organised along disciplinary lines (20 in all). Panel members were appointed by the CSIC upon nominations by the European Science Foundation (ESF) and the European Molecular Biology Organization (EMBO).

The evaluation took into consideration the whole capacity of each institution and assessment dimensions were:

- Mission, vision and strategy of institutes
- Research quality
- Manpower and resources
- Present and future organisation of institutes and research perspectives.

Discussions [Session A3 and B3]

In both sessions, comments and discussions following the presentations and at the end of the sessions echoed the speakers' views that the evaluation of research institutions is essentially different from the evaluation of individual research grant applications. Therefore criteria used for assessing institutions have to be different, more diversified, than for assessing research grant applications.

The following critical issues were identified:

(1) The broader mission of research institutions Scientific institutions, be they universities or research institutes, have a broader mission than just research. Assessments of their performance should therefore take into account the quality of projects (excellence) but should also consider societal needs.

In the case of universities there is a need for regional/societal access to high quality higher education. A possible way to accommodate potentially conflicting needs could be a combination of different streams of funding ensuring support for excellence in research and for other types of performance important for society (teaching, continuing education, technol-

ogy transfer). For example, the Linnaeus Grants in Sweden are only one of several streams of funding for universities. They are awarded solely on the research excellence principle while other funding sources ensure regional coverage of higher education institutions.

In the case of research institutes, other missions such as providing access to research infrastructures and services e.g. IFREMER, should also be taken into consideration.

The scope of the evaluation exercise goes beyond the assessment of scientific performance. Usually the strategic visions of the institutes, their management and the quality of their services are also evaluated at the same time. Therefore the methods of their assessment should be made clear. In this context Session A3 gave the Standard Evaluation Protocol 2003-09 for Public Research Organisations² as an example of an elaborate evaluation scheme, which takes into account the various missions of publicly funded research institutions. This protocol outlines a new system of evaluation of research institutions which is being used in publicly funded research in the Netherlands. The universities, the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO) agreed on the use of this protocol in any research evaluations they will undertake or commission.

(2) The long-term perspective

Research institutions, unlike most individual research projects, operate on a long-term basis. Their reviews should take into consideration the need for stability of funding for the institutions both to ensure the development of a given research area and to secure the professional careers of researchers. As a way of complying with this vision, the suggestion was made to decouple results of evaluation with immediate funding decisions, allowing a 'buffer' between the two processes.

(3) The scope of evaluation

Distinct from the evaluation of research grant applications that concentrate on the scientific quality of a proposal, the evaluation of institutions must find a balance between looking back and presenting daring new activities. In the case of Linnaeus Grants, universities were competing on the basis of the attained scientific quality (state-of-the-art) and potential for development (future strategy). The same approach was used for assessing strategic plans of CSIC institutes where future strategies were based on past achievements. In the INSERM team evaluation exercise, bibliometric indicators were used as a complementary tool in the assessment of up to date achievements of research teams, which in turn were treated as a prognosis of

future achievements. The assessments are thus facing a difficult task: to be both retrospective and prospective.

(4) Internationalisation of review panels

Very often, because a national system of institutions which undergoes evaluation involves a large part of the research community of a given country, there is a need to widely use international peers. The examples given in the sessions showed that, increasingly, international peers were being invited to be part of the review panels: Linnaeus Grants where only foreign peers were assessing proposals and making decisions, CSIC evaluation where a substantial number of reviewers from outside Spain were invited, the evaluation of the Italian National Institute of Nuclear Physics where an international evaluation committee was appointed.

(5) Information technology

Discussing the use of information technology to support the process of evaluating research institutions, it was suggested that information systems and databases give access to a very wide, international pool of experts and, helping to identify the best expertise for a given case, enhance the quality of the peer review.

Another area in which information technology may help in the evaluation of research institutions is its potential to reduce the burden of evaluation exercises put onto researchers. Creating data repositories was suggested as a way to alleviate some of the administrative burden.

(6) The timelines and improvement cycles

Choosing the periodicity thoughtfully would not only reduce the burden on researchers of the evaluated institutions but would also help getting feedback on improvement made and suggestions for changes in due time.

3.4. Science Policy Context for Excellent Peer Review [Session C1]

Session C1

Chair: **Dieter Imboden**, *Swiss National Science Foundation*

Ian Halliday, *European Science Foundation*, Does peer review exist?

Chris Caswill, *University of Exeter*, Reasons for continuity and pressures for change – the scope for fresh thinking on peer review

Klaus Zinoecker and Michael Stampfer, *Vienna Science and Technology Fund*, Peer review and beyond?

Rapporteur: **Mariana Resnicoff** (ESF)

2. <http://www.knaw.nl/publicaties/pdf/90000091.pdf>

3. Parallel Sessions



From the left: K. Zinoecker, D. Imboden, I. Halliday, Ch. Caswill

The diversity of the approaches with regard to the practical organisation of peer review is matched by the wide range of the areas in which it is used. In the context of research institutions the peer review is used to judge the promise of a research grant/grantee; to assess the progress of on-going research programmes and to retrospectively rate the quality of the outputs of funded research.

This session, chaired by Professor Dieter Imboden, President of the Swiss National Science Foundation, tried to stand back and look at the context in which it operates:

- Are strategic visions being sufficiently taken into account by current peer review mechanisms?
- What are the challenges it faces, where do they come from and are they being met?
- Are there possible alternatives in allocating research funding?

Professor Halliday, President of the European Science Foundation, challenged in his presentation the presumption of peer review (that any reasonable panel will come out with the same result with respect to excellence), and argued that the results are actually dependent on prior assumptions which should be made transparent before the review. The importance of supporting innovative frontier science and the difficulties encountered in the current climate of tight budgets and timelines were discussed. The use of a 'roadmap' and financial spreadsheets was presented as a powerful aid for peer review panels in making strategic decisions. Peer review needs a science policy context and iterative processes to achieve excellence; scientific quality and science policy have to be coupled.

Professor Halliday called for innovative research to be embedded in the system rather than launching new schemes to support this kind of research.

The challenge for the future is to overcome the risk-averse funding culture in Europe and the so called 'tyranny of peer review': decisions based on past investments, big laboratories and/or conservatism.

The presentation of **Professor Caswill**, former Director of the Economic and Social Research Council UK, discussed the diversity of peer evaluation activity and interest, looking at the underlying issues of utility, knowledge production and knowledge transfer, and the significance of new techniques and IT tools in future developments and their impact on society and the economy.

Innovative research is confronted with bias, conservatism and politics; from an economist's perspective, peer review can be seen as an economic market, where there is a struggle for power: scientific authority versus social capital. It is imperative to consider innovation as the key driver of economic performance and to include the socio-economic impact of innovation in the evaluation criteria. Knowledge transfer is currently a less visible driver but it should be considered a main evaluation output.

New techniques and the so called 'ICT push' supply additional pressures to the peer review system. These are 'healthy' pressures because they lead to improvements and increase the quality of the peer review process. The evaluation industry is pushing for more evaluations, including analyses of cost-benefit ratios, surveys and bibliometrics; unfortunately, evaluation is too often reduced to mere techniques rather than the use that can be made of it.

The challenges for the future are the understanding of impact in terms of utility, knowledge production and knowledge transfer, and the involvement of new stakeholders in the decision-making process to assess the impact of the research in society and the economy.

Mr. Zinoecker discussed and assessed several strategies to modify or partly replace peer review in scientific selection processes.

He noted that in cases in which peer review is used in the evaluation exercise, for example in the evaluation of research institutions; a combination with other methods (such as professional evaluators) may be beneficial to the exercise.

Discussing peer review in assessment of research grants he listed several strategies. Among them are the following:

- Focal randomisation mechanism (FRM) which suggests funding the projects unanimously ranked at the top by all reviewers, rejecting those deemed valueless while funding the rest on a random basis.
- The Dutch Innovation Voucher Programme which aims at increasing the interaction between Small and Medium Enterprises (SMEs). A voucher (worth €7 500) can be spent by an SME to commission research at a public institution. Vouchers are allocated randomly by means of a lottery.

Mr. Zinoecker concluded that peer review is not replaceable by any of the suggested strategies but that a search for alternatives or mechanisms to improve it should be continued.

3.5. Language, Metrics, Impact: Cultures of Peer Review in the Humanities [Session C2]

Session C2

Chair: **James Herbert**, *Centre for Research in the Arts, Social Sciences, and Humanities, Cambridge University*

Anton Nederhof, *Leiden University*: Quantitative assessment of research performance, peer review and the role of language in the humanities

Emmanuela Reale, *CNR, CERIS*, Peer review and the humanities – strengths and weaknesses for research evaluation purposes

Michael Worton, *University College London*, Of models and metrics – the UK debate on assessing humanities research

Professor Jakob de Haan, *University of Groningen*, How reliable are journal impact scores as quality indicator? Evidence for the field of economics

Rapporteur: **Rüdiger Klein** (ESF)

In an environment in which evidence-based decision making is quickly gaining currency, the adequate assessment of humanities research (ex-ante as well as ex-post) is problematic. The session discussed the role of peer review as the necessary basis for the construction of adequate tools needed to evenly access and assess research output in the humanities.

Dr Nederhof (the Centre for Science and Technology Studies, Leiden), provided evidence for the unsatisfactory state of metrics for the humanities, where bibliometric tools do not, at present, allow the capture of diverse publishing behaviour of humanities scholars. The existing tools such as the widely used ISI database of Thomson Scientific and its use to perform impact factor analysis can be useful only for few areas of the humanities, and only where internationalisation has led to a prevalence of publications in English (and in Anglo-Saxon) journals. However, the three audiences of humanities research – international and national scholars, as well as the non-academic, educated public – make for a characteristic distribution of publication in international journals, other journals and monographs, often in different languages. Notably in ‘identity’-linked disciplines, such as literature or history, publishing in national languages (or, indeed, object languages) prevails, while in more internationalised fields (such as large parts of linguistics) this is no longer the case. Generally, journal publications are less relevant than in the other sciences, even though publishing behaviour seems to be changing. It is expected that the availability of the European Reference Index for the Humanities (ERIH)

3. Parallel Sessions



From the left: A.J. Nederhof, E. Reale, J. de Haan, M. Worton

will present a much better basis on which to construct appropriate metrics; it is recommended, however, that books be included in the future work-plan.

Dr Reale (Istituto di ricerche sull'Impresa e lo Sviluppo CERIS del CNR) tested in her presentation the reliability of the peer review process against metrics calculated on the basis of records in the ISI database. The test case was the Italian three-year Research Assessment Exercise. Journal impact factors were shown not to be good indicators for the quality of articles, as assessed during the peer review process. A remarkable sign for the independence of the peer review judgment was, on the other hand, the appreciation of interdisciplinary work (over traditional approaches), notably visible in the panel covering philosophy-psychology-pedagogy and history. Conversely, it emerged that bibliometrics, where they exist (e.g. in economics), do shape peer review behaviour, and are likely to also shape publication behaviour. In this process – linked as it is to internationalisation – national-language publications are slowly losing ground.

Professor Worton (University College London) reported on the deliberations of the joint HEFCE (Higher Education Funding Council for England) and The Arts and Humanities Research Council (AHRC) expert group on research metrics (2006), over which he presides. The group's work rests on the assumption that an assessment system that takes into account disciplinary specificities as much as changing research landscapes needs to be developed. Only a portfolio of metrics will be able to produce relevant information about the entire research process (input-activity-output-outcome). He argued that there is no fundamental

difference between scientific, technological, engineering and medical disciplines and the humanities for the purposes of the research assessment. Public access reference tools currently developed outside the UK, such as the emerging European Reference Index for the Humanities developed by ESF on behalf of its Member Organisations, may contribute to setting international benchmarks.

Professor de Haan (Economics, University of Groningen) provided evidence from the field of economics for the question of how reliable journal impact scores can be as a quality indicator. He criticised the 'admission' policy by the currently dominating SSCI (Social Science Citation Index by ISI Thomson Scientific), and argued that the criteria do not do justice to the research and publication process, while at the same time laying authors open to pressure from publishers that may lead to a distortion of citation figures. Professor de Haan compared the outcome of impact factor analyses following ISI with other indicators and came to the conclusion that the ISI system is too closed to be useful for quality assessments. He advised humanities against accepting any suggestions that simple impact factor metrics would provide a reliable tool for assessing research quality.

Discussions [Session C2]

The discussion ranged over a variety of topics, from the likely impact of open-access depositories to questions relating more directly to the need to study the social processes underlying peer review as much as gatekeeper behaviour in journals. It was emphasised that the focus on journals in current metrics follows the practice in the natural sciences, and that this may be necessary in terms of creating a level playing field. It also emerged, however, that – if given an opportunity – natural scientists would prefer more subtle metrics such as are now being developed for the humanities, with their stronger role for discursive peer review.

Humanities research must be seen in the context of its three audiences – international and national scholarly communities, as well as the educated, non-academic public – and its influence on identity-generating processes must not be underestimated. Therefore any harmonising process across Europe must aim at capturing linguistic diversity, and diverging publication patterns in the different sub-disciplines. This can only be done if it is understood that metrics need to be supported by peer review, and vice versa.

One needs to proceed with an awareness of the existing mismatch between the available bibliometric tools and the nature of humanities research and publication patterns. There was an agreement that it would be desirable to develop metrics that would support peer review in a more accurate fashion than the currently used tools.

The panel came to the conclusion that the first step would have to include exercises that aim at mapping and accessing (initially) journal-based, multilingual European humanities research output of the highest quality such as the ESF project the European Reference Index for the Humanities (ERIH) before any further steps can be taken.

European Reference Index for the Humanities (ERIH)

Humanities research in Europe is rich in lively national linguistic and intellectual traditions. They all find their expression in scholarly publications. The new transnational mobility of researchers and the often transdisciplinary nature of contemporary science require that humanities researchers position themselves in changing international contexts. Researchers and institutions need a tool that helps to evenly access the scientific quality of humanities research output, irrespective of disciplinary and linguistic boundaries.

The ESF project, the European Reference Index for the Humanities, will provide such a service: a reference index of top journals in 15 areas of the humanities. In a peer review based, Europe-wide process, 15 expert panels aggregate and sift input received from funding agencies, subject associations and specialist research centres from across the continent and beyond.

It is envisaged that the ERIH will eventually be extended to include book-form publications and non-traditional formats, so that it can operate as a backbone to a fully-fledged research information system for the humanities.

The ERIH has been sponsored in 2005/06 by ESF and the European Commission ERA-NET project Humanities in the European Research Area (HERA).

4. Concluding Session

Chair: **Peter Nijkamp**

Panel Members:

Bertil Andersson, Hans Gudmundsson, James Herbert, Gabor Makara, John Marks, Pär Omling, Motoyuki Ono, Václav Pačes, Stephen J. Simpson, Josef Syka, Raimo Väyrynen

Rapporteur: **Eva Olde Monnikhof**

The concluding session aimed at capturing the various contributions made in the parallel sessions and providing an opportunity to discuss future common activities.

The Chair asked the panel members to report back on the various sessions, taking into consideration the three key questions the conference addressed:

- Is peer review in the present form able to identify the best and most innovative frontier science and how might it be improved?
- What is the best way to harmonise the peer review process and how can new methods and IT tools contribute to it?
- What are the major societal, cultural and ethical challenges of future peer review processes and how could they be incorporated?

Is peer review in the present form able to identify the best and most innovative frontier science and how might it be improved?

The scientific landscape has changed under the influence of the European Research Area, the overall

globalisation and the move in research towards multi-disciplinarity. Nevertheless the peer review system has remained virtually the same. Since it was concluded that there are no real alternatives to peer review there is a need to improve the system and look at the system in a different way. In most of the discussions it was recognised that the system is not flawless. There is a need to pay more attention to innovative frontier science which is often high-risk research. Several highly important innovative developments in research may be rejected because of conservative views within the system. It has been mentioned that there is no place for conservatism in the peer review system.

Specificities in the use of peer review for evaluation of institutions were pointed out. In reviewing the performance of institutions, not only their attained scientific quality but also the potential to realise strategic development must be considered. Furthermore, although undoubtedly scientific excellence is the most important goal of the institution also societal concerns for equity and fairness must be taken into account when evaluating science institutions.

What is the best way to harmonise the peer review process and how can new methods and IT tools contribute to it?

Instruments such as EUROCORES and EURYI have contributed to the harmonisation of the peer review process at the European level. The close cooperation between the Scandinavian countries has also led to more harmonisation of the process. Nevertheless lack of harmonisation



Closing of the conference (from the left: S.J. Simpson, J. Herbert, G.B. Makara, M. Ono, P. Omling, H.K. Gudmundsson, B. Andersson, J. Marks and J. Syka)

remains a difficult hurdle to overcome because of differences in the scientific traditions and research systems between various countries. However, the increasing use of international referees creates a need for more harmonisation (e.g. using English as the language of science). Efforts are needed to try to harmonise procedures and practices both in peer review and in the evaluation of institutions.

Another area which may need further exploration is the education of reviewers in the tasks they are expected to do and the professionalisation of science management. The potential increase in R&D funding in Europe will put new burdens on officers of the funding agencies and research councils. Their high professionalism will play an essential role in the future quality of the peer review process in Europe.

Guidelines for a minimum set of standards should be developed (e.g. in EuroHORCs and/or ESF) in order to 'find one another' in a European procedure and ensure trust and liability. Grant agencies, research councils and other bodies distributing money for research on the basis of peer review in Europe should collaborate more closely, e.g. by sharing their databases of reviewers. Furthermore, organisations should try to use each other's review panels. A most desirable approach, mentioned several times, was the creation of a common pool of reviewers in Europe. As an initiative to discuss possible cooperation between European research organisations the ESF Member Forum on Peer Review was announced at the conference.

ESF Member Organisation Forum on Peer Review

For ESF Member Organisations, peer review and grant-awarding procedures are the keys to the quality of their performance and to their image in the scientific community. However, with changes to the ways research is organised and funded, new challenges and requirements for peer review arise. This conference was also a launching event for the ESF Member Organisation Forum on Peer Review.

The forum will serve as an opportunity to exchange experiences and develop best practices, also those resulting from the large number of ERA-NETs dealing with comparisons of peer review systems in which ESF Member Organisations are involved. One purpose of the forum is to bring together these experiences including the expertise developed by ESF in the framework of EURYI and EUROCORES. Interested Member Organisations held their first meeting during the conference in Prague. The forum will continue into 2007.

Activities and outcomes of the forum will include working groups, workshops, preparation of reports, or best practice guidelines as well as collaboration and resource-sharing agreements.

For further information please contact Dr Nina Hoffmann (nhoffman@esf.org), who coordinates within the ESF the activities of this ESF Member Organisation Forum.



4. Concluding Session

Electronic tools must be used widely, especially in cases where referees come from all over the world. They can also be used in the effort to coordinate reviewers' inputs, set up meetings, sharing databases, etc. It is expected that many of the activities of peer reviewers connected with travelling will be replaced in the future by electronic platforms such as teleconferences, video-enhanced discussions and secure Internet discussion boards.

What are the major societal, cultural and ethical challenges of future peer review processes and how could they be incorporated?

It was recognised that each country has its own specificities also regarding peer review. In some countries it is common practice that a lot of emphasis is put on the status of the applicant while in other countries the focus is more on the application itself. The question of remuneration of referees was mentioned in several contributions. No standard approach exists among funding agencies.

Another issue that came up during the discussion is the difference between domestic and foreign reviewers. In some countries domestic reviewers are seen as 'stricter', in others their view seems 'more soft' than that of the foreign reviewers.

In discussions about the 'best reviewers' it was pointed out that 'best reviewers' must not only be excellent in their field but also open to new ideas and tolerant of scientific approaches which they may not necessarily share.

Also the issue concerning conflict of interest (or perceived conflict of interest) was touched upon. Avoiding any conflict of interest is the cornerstone of any effective peer review process. Therefore a set of guidelines or protocol to tackle them should be developed. There are several guidelines already existing in different funding agencies and research councils that can be followed.

In addition, the misjudgement of referees was mentioned. It is not possible to avoid this in all cases. Referees are people and people tend to make mistakes. In efforts to assure a fairer judgment of proposals, some agencies ask the referees to self-rate their expertise. Examples of software being used to assign proposals to referees on the basis of key words were mentioned as well as the introduction of possibilities for 'rebuttal' from applicants.

Moreover it was noted that there should be more attention paid to cultural differences in disciplines. Special interest was given to problems of peer review in the humanities. A recommendation was made to further develop tools to measure the impact of research in humanities in Europe such as the ESF European Reference Index for the Humanities.

Recently, serious ethical problems have appeared in the publication of results of research in scientific journals.

Several cases of fraud, fabrication and falsification of data have led publishers, editorial boards and ethical committees of scientific institutions to strengthen the peer review process. Yet the main responsibility for investigating allegations lies with the institutions and funding agencies that pay for the work.

In a number of research fields, notably in biomedicine, science journals have started to experiment with an open peer review system. Instead of using traditional blind peer review, they publish reviewer's reports (thus making the names of reviewers public).

All the abovementioned aspects and problems of the contemporary peer review need a strategic evaluation at a more conceptual level, not of a specific mechanism, but of the underlying principles and ideas.

Continuously monitoring the peer review system on which the quality of science so heavily depends will contribute to its improvement. The conference represented an effort in this direction. It helped to identify the issues faced by peer review systems today, to raise the awareness of potential critical pitfalls and to exchange experiences on how these are tackled by different organisations.

Appendices

Appendix 1

Conference Programme

Thursday, 12 October 2007

9.00 – 13.00: Plenary Session:

Role of the peer review in the new scientific landscape; the notion of excellence
Chair: **Ian Halliday**

G. Stroud: Evaluation rules and procedures under the EU framework programme for research and technological development: the current state of play and future developments

T. Scarpa: NIH peer review: challenges and opportunities

P. Campbell: Pressures on peer review at Nature journals

K. Mehlhorn: Peer review in the electronic age
S. Hornbostel: Peer review: in the core healthy or chronically ill?

J. O'Reilly: Risk, adventure and the tyranny of peer review

14.30 – 16.00: First series of parallel sessions

Session A1: Assessment and selection of proposals: methodologies and processes
Chair: **Václav Pačes**

H. Bruun: Evaluating interdisciplinary research proposals: the case of the Academy of Finland

O.-K. Kwon and J. Lim: Role of peer review system in the evaluation of the interdisciplinary research proposals at the KOSEF, Korea

P. Chráska and J. Syka: Peer Review in the Czech Science Foundation - experiences and challenges

N. Yetis, O. Cebeci, A.Z. Oruç: Effective evaluation and funding of research projects (TÜBITAK)

Session A2: European level peer review; peer review of networking activities: international
Chair: **Pär Omling**

L. Grivell: Peer review at EMBO

F. Fedi: Assessment, monitoring and final evaluation of COST actions

L. Allen: Looking for landmarks: the importance of post-award peer review

Session A3: Evaluation of institutions
Chair: **Gábor Makara**

R. van Duinen: Review of research institutes and programmes; experience and reflections

L. Devenon and J.-Y. Perrot: Evaluation of research units at IFREMER

A. Bertin: The evaluation experience of the Italian National Institute for Nuclear Physics: state-of art and open problems

Thibaut Lery: Avoiding double peer review in the provision of scarce infrastructure

16.30 – 18.00: Second series of parallel sessions

Session B1: Assessment and selection of proposals: global challenges and national practices
Chair: **John Marks**

Ch. Schneider: The worst form of deciding on grant proposals except all others - an administrator's view on peer review

M. He: Discussion on peer review of grant applications in NSFC

D.-P. Min and D.S. Kang: Assessment system for grant proposals in Korea

Session B2: European level peer review; peer review of networking activities: collaboration of national research funding agencies
Chair: **Bertil Andersson**

Invited Panel members: **R. Väyrynen, I. Abram, R. Novak, B. Scholz**

Session B3: Evaluation of institutions cont.
Chair: **Hans K. Gudmundsson**

P. Omling: Linnaeus Grants – university funding in national competitions

N. Haeffner-Cavaillon: Peer review and bibliometrics

R. Rodrigo: The role of institutional evaluation in the CSIC strategic planning

Friday, 13 October 2007

8.30 – 10.10: Third series of parallel sessions

Session C1: Science policy context for excellent peer review
Chair: **Dieter Imboden**

I. Halliday: Does peer review exist?

Ch. Caswill: Reasons for continuity and pressures for change: the scope for fresh thinking on peer review

K. Zinoecker, M. Stampfer: Peer review - and beyond?

Session C2: Language, metrics, impact – cultures of peer review in the Humanities
Chair: **James Herbert**

A. Nederhof: Quantitative assessment of research performance, peer review, and the role of language in the humanities

E. Reale: Peer review and humanities: strength and weaknesses for institutional evaluation purposes

M. Worton: Of models and metrics: the UK debate on assessing Humanities research

J. de Haan: How reliable are journal impact scores as quality indicator? Evidence for the field of economics

10.30 – 13.00: Plenary session:

Global Contexts of Peer Review
Chair: **Josef Syka**

A. L. Bement: The reason and reach of merit review at the U.S. National Science Foundation

B. Nordén: Peer review and young research careers : EURYI scheme

D. Estève: Peer review system of the European Research Council

M. Ono: JSPS's application review system - present and future

S. Simpson: Exploring the limits of peer review in science publishing

14.30 - 16.00: Summary of parallel sessions and general discussion

Chair: **Peter Nijkamp**

Appendix 2

Conference Participants

Arthur Aamodt	The Research Council of Norway
Izo Abram	CNRS
Rachid Adghoughi	ESF
Jüri Allik	Estonian Science Foundation
Liz Allen*	The Wellcome Trust
Bertil Andersson*	ESF
Bernard Avril	ESF
Ge Bai	National Natural Science Foundation of China
Huba Bartos	OTKA
Raymond Bausch	Fonds National de la Recherche Luxembourg
Arden L. Bement, Jr.*	U.S. National Science Foundation
Antonio Bertin*	National Institute for Nuclear Physics
Hakan Billing	Swedish Research Council
Sofie Bjorling	Swedish Research Council
Imre Boros	University of Szeged
Andrew Bourne	Engineering and Physical Sciences Research Council (EPSRC)
Henrik Bruun*	IASM, Turku
Philip Campbell*	Nature
Chris Caswill*	University of Exeter
Alan Cross	European Commission
Boryana Damyanova	Bulgarian Academy of Sciences
Jakob de Haan*	ESF, Standing Committee Social Sciences/University of Groningen
Franci Demsar	Slovenia Research Agency
Daphne den Hollander	Netherlands Organisation for Scientific Research (NWO)
Jean-Luc Devenon*	IFREMER
Jan Dobeš	NuPECC
Michel Dodet	Institut National de la Recherche Agronomique
James Dutt	Institute of Experimental Medicine ASCR
Jüri Engelbrecht	Estonian Academy of Sciences
Daniel Esteve*	Scientific Council of the ERC
Francesco Fedi*	COST
Hans Fink	Danish Research Agency
Otakar Fojt	British Embassy in Prague
Philippe Galiay	European Commission
Bohuslav Gas	Charles University in Prague
Victoria Gill	ESF
Claude Graillot-Gak	INSERM
Les Grivell*	EMBO
Hans K. Gudmundsson*	RANNÍS, The Icelandic Centre for Research
Vladimir Haasz	Council of Czech Universities
Barbara Haberl	Austrian Academy of Sciences
Nicole Haeffner-Cavaillon*	INSERM
Ian Halliday*	ESF
Lise Hansen	The European Court of Auditors
Minghong He*	National Natural Science Foundation of China
Berit. L. Heitmann	Institute of Preventive Medicine, Research Unit for Dietary Studies
Jean-Pierre Henriet*	Ghent University - Department of Geology and Soil Science
James Herbert*	NEH (ret) / CRASSH
Wolfgang Höritsch	Oesterreichische Nationalbank- Jubiläumsfonds
Stefan Hornbostel*	Institute for Research Information
Gudmund Host	The Research Council of Norway

* Speakers / Chairs of session

Jan Hrušák	Academy of Sciences of the Czech Republic
Huai Chen	National Natural Science Foundation of China
Pavel Chráska*	Czech Science Foundation
Dieter Imboden*	Swiss National Science Foundation
Chanan Itai	Ben Gurion University
Arja Kallio	ESF
Nina Kancewicz-Hoffman	ESF
Dong-Seob Kang*	Korea Research Foundation
Ruediger Klein	ESF
Elisabeth Kokkelkoren	FNRS
Laszlo Kollar	OTKA
Miroslava Kopicová,	Czech Minister of Education, Youth and Sports
Dimitros Kyriakidis	National Hellenic Research Foundation
Lars Lading	Danish Agency for Science Technology and Innovation
Vladimir Lapshin	Russian Foundation for Basic Research
Jacqueline Lecourtier	ANR
Thibaut Lery	Dublin Institute for Advanced Studies
Jong Geon Lim*	Korea Science and Engineering Foundation (KOSEF)
Pero Lucin	The National Foundation for Science (Croatia)
Gábor Makara*	OTKA
John Marks*	ESF
Martín Martínez-Ripoll	Spanish Council for Scientific Research (CSIC)
Kurt Mehlhorn*	Max-Planck-Society
Sandra Milovanovic	National Foundation for Science (Croatia)
Luule Mizera	Estonian Science Foundation
Jürgen Mlynek	Helmholtz Society
Elisabeth Monard	Research Foundation Flanders (FWO)
Alexis-Michel Mugabushaka	ESF
Annette Moth Wiklund	Swedish Research Council
Ton Nederhof*	CWTS University Leiden
Vladimír Nekvasil	Institute of Physics, ASCR
Ebba Nexo	Danish Independent Research Council
Peter Nijkamp*	Netherlands Organisation for Scientific Research (NWO)
Bengt Nordén*	Chalmers University of Technology
Rudolf Novak*	FWF, Austrian Science Fund
Jussi Nuorteva	Finish Academies of Science and Letters
John O'Reilly*	EPSRC
Bo Ohngren	Swedish Research Council
Kazuhisa Okamoto	Japan Society for the Promotion of Science
Eva Charlotte Olde Monnikhof	NWO
Pär Omling*	Swedish Research Council
Motoyuki Ono	Japan Society for the Promotion of Science
Yavuz Oruç*	Scientific and Technological Research Council (TUBITAK)
Gerald Owenson	UK Biotechnology and Biological Sciences Research Council
Václav Pačes*	Academy of Sciences of the Czech Republic
Veronika Palečková	Czech Science Foundation
Petr Ráb	Academy Council AS CR
Emanuela Reale*	CERIS CNR
Martin Reinhart	Science Studies Program, University Basel
Mariana Resnicoff	ESF
Jong Hyun Rhie	Korea Science and Engineering Foundation (KOSEF)

* Speakers / Chairs of session

Appendix 2

Gareth Roberts	The European Court of Auditors
Daniele Rod Wiesner	SwissCore - Swiss National Science Foundation
Rafael Rodrigo*	CSIC
Giuseppe Roffi	CNR/EuroHORCs
Peter Samuely	Slovak Academy of Sciences
Kate Sapir	Israel Science Foundation
Toni Scarpa*	Center for Scientific Review, NIH
Joao Sentieiro	FCT
Elena Shkarina	Russian Foundation for Basic Research
Beate Scholz*	DFG
Christoph Schneider*	DFG
Stephen J. Simpson	Science Magazine
Meelis Sirendi	Estonian Science Foundation
Daniel Sirtes	Science Studies Program, University Basel
James Slak	Slovenia Research Agency
John H. Smith	European University Association
Kateřina Sobotková	Czech Science Foundation
Jan Sokol	Czech Science Foundation
Michael Stampfer*	WWTF Vienna Science and Technology Fund
Lars Stemmerik	Geological Survey of Denmark and Greenland
Graham Stroud*	European Commission
Mark Suskin	U.S. National Science Foundation
Josef Syka*	Czech Science Foundation
Kumiko Tansho	Japan Society for the Promotion of Science
Siri Tonseth	The Research Council of Norway
Jan Uhlř	Czech Science Foundation
Rein Vaikmāe	Tallinn University of Technology
Jacco Van den Heuvel	Royal Netherlands Academy of Arts and Sciences (KNAW)
Monique van Donzel	ESF
Reinder van Duinen*	ESF
František Vaníček	Czech Science Foundation
Raimo Väyrynen*	Academy of Finland
Patricia Vogel	NWO
Pavel Vlasák	Academy of Sciences of the Czech Republic
Jean Frank Wagner	EuroHORCs/FNR Luxembourg
Jean-Bernard Weber	Swiss National Science Foundation
Marcel Weber	Science Studies Program, University Basel
Torsten Wiesel	The Rockefeller University
Michael Worton*	UCL (University College London)
Klaus Zinoecker*	WWTF Vienna Science and Technology Fund
Petr Zuna	Czech Technical University in Prague

* Speakers / Chairs of session

Contacts

Daphne den Hollander	EuroHORCs	Hollander@NWO.NL
Nina Kancewicz-Hoffman	European Science Foundation	nhoffman@esf.org
Veronika Palečková	Czech Science Foundation	veronika.paleckova@gacr.cas.cz

Published by the European Science Foundation.
March 2007.

Suggested citation:

European Science Foundation: *Peer review – Its present and future state, Prague, 12-13 October 2006. Conference report.*
ESF, Strasbourg, March 2007

This report incorporates the opinions expressed at the conference. They do not necessarily represent the views of the European Science Foundation or its Member Organisations.

Cover: Corbis

Photo credit: The Czech Science Foundation (GAČR)

Printing: IREG Strasbourg

Report Editing: Alexis-Michel Mugabushaka

ISBN: 2-912049-60-1

