ACUMEN
academic careers understood through measurement and norms

FP7 Grant Agreement | 266632

Deliverable No and Title	**1.3. Peer Review Practices**
Dissemination Level	PU (public)
Work Package	1
Version	3
Release Date	15.10.2012
Author(s)	Ülle Must, Oskar Otsus, Marcus Denton, Arto Mustajoki, Taavi Tiirik

Introduction

Peer Review (PRev) is among the oldest certification practices in science and was designed to prevent poor research from taking place. There is overall agreement that PRev is the most solid method for the evaluation of scientific quality. Since PRev spans the boundaries of several societal communities, science and policy, research and practice, academia and bureaucracy, public and private, the purposes and meaning of this process may be understood differently across the communities. In Europe, internationally competitive research activities take place in large superstructures as well as in small, insufficiently funded university departments; research can be publicly or privately funded; the purpose may be applied research often with a focus on the needs of regional industry, or purely 'blue-sky' research.

In current report we focused mainly in on PRev of grant applications, the analysis has been carried out on the basis of **PRev related literature analysis** (Thomson Reuters, Union Library Catalogues, Google Scholar, and reports of selected research funding organisations).

The outcome of this task is an open dataset of PRev related documents archimedes.ee/acumen), and a review of PRev related documents. **The questionnaire survey** was conducted with the aim to receive more insight into researchers' attitudes towards the ways in which quality, success, excellence and impact of scientific production are measured and evaluated, and get suggestions about how the current PRev system should be improved or modified. In the questionnaire, the focus of interest was on researchers' views and experience as either a project applicant or reviewer.

Data and methods

The documentary analysis was conducted on the basis of published papers and monographs on the process of Peer Reviewing (available from Thomson Reuters, Union Library Catalogues, Google Scholar), and the survey reports of the research funding organisations. Articles and texts were identified through targeted searches, using terms including "grant peer review", "peer review AND funding" , after which manual cleaning was performed. The main purpose was to select documents which contain materials about research on Peer Review, and especially on PRev of grant applications. In total 380 documents were selected. The report is divided into six blocks:1. Peer Review at the European Union level; 2. European Science Foundation initiatives; 3) National Peer Review studies; 4) Pros and cons of Peer Review Process versus Metrics; 5. Peer Review practices (PRP); 6. Peer Review biases.

A structured web based (LimeSurvey) questionnaire was prepared for the survey. The survey was divided into four blocks:

1. General information (fields of science and technology, country of affiliation, citizenship, gender, the profile as researcher);

1.2. Experience in the PRev process (as applicant or reviewer);

1.3. Respondents' experience as an applicant (how many applications they have made, how successful on an average have they been, being informed of the reasons for the rejection, the reasons for the rejection);

1.4. Respondents' experience as a reviewer (how many reviews they have made, have they refused review offers, what were the main reasons for refusal, have respondents been informed about the end results of the application reviewed by them, would this feedback be necessary);

2. S&T indicators (the most appropriate indicators in assessing the work of researchers, should the various indicators be weighed differently, is there a need to have different weightings for the various indicators in different subject areas, is there a need to have various indicators or different weighing for the various indicators at different career stages);

3. Criticism of PRev (effects of different bias on the assessment of the applications, the most important issues to consider for optimising fairness and objectivity in the evaluation process, what are the most essential criteria for a good reviewer, to what extent information about reviewers should be available to the applicants);
4. The future of the peer review system.

The survey was open for two months (November – December 2011). The information and call to participate in the survey was sent out via different communication channels. In total, 2114 respondents answered.

1. Literature Survey

PRev as one of the methods of research evaluation has been subjected to a variety of evaluative activities over the past 30 years. As we see from Figure 1, the number of papers on the PRev issue has dramatically increased, the same trend is observed with respect to collaboration. While in the 80s the proportion of papers written in collaboration was 18.4%, it is 53.7% in the 21st century.

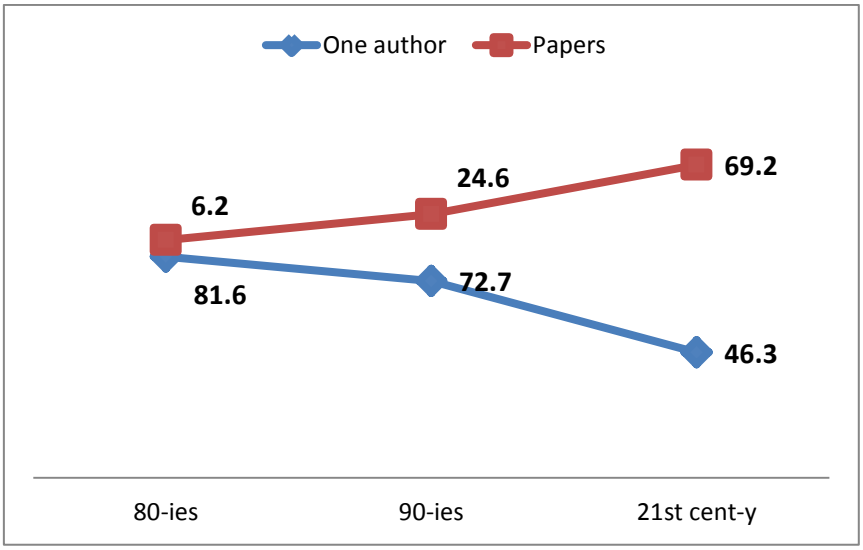

Figure 1. The proportion of single authored papers and total papers by time period

Papers vary by fields, about one third of papers are written in medical sciences. Here we have to take into account that in medicine the term of PRev means not only research evaluation but also practical reviewing of the work. In the current survey we deal only with papers related to PRev in research. Social sciences and natural science are represented on the same level – about a quarter of total amount of written papers.

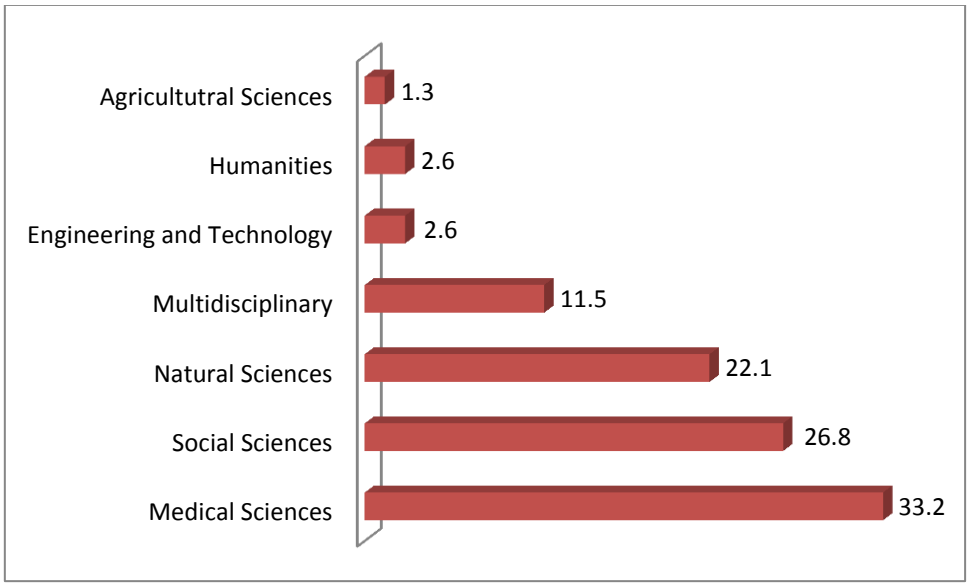

Figure 2. The proportion of papers by fields (in %)

In Europe, internationally competitive research activities take place in large superstructures as well as in small, insufficiently funded university departments; research can be publicly or privately funded; the purpose may be applied research often with a focus on the needs of regional industry, or purely 'blue-sky' research. Research policy and funding schemes vary considerably in the member states, and even though PRev is widely accepted there is no agreed common standard for quality assurance. In 1995 Luke Georghiou stated in his paper (Georghiou, 1995) that research evaluation experience in European national science and technology systems and institutional reforms have created a demand for a new kind of evaluation.
Although peer review activities have been the focus of research funding bodies for 30 years, it has truly begun as systematic co-operation in the 21st century. Such supranational bodies as European Science Foundation, Nordic Council, and European Commission are standing in the forefront.

Peer Review at the European Union level

The European Commission has developed and implemented the evaluation of research since the end of 1970s when the Evaluation Unit of the Commission was created. Mid-term and ex post evaluations are organised with the help of independent experts. The panel members must be independent from the programme directors and the Commission and well known experts in the field (Bobe and Viala, 1997; Massimo, 1997). The Commission keeps a database of potential evaluators which is open for anyone to register (https://cordis.europa.eu/emmfp7/index.cfm).
The criteria for the panel members are as follows: a) a good coverage of scientific knowledge relevant to the programme; b) interest in evaluation methodology; c) familiarity with socio-economic impact. (Bobe and Viala, 1997)

Expert Groups Reports

In the 2010, two important papers were published by European Research Area Committee (ERAC). One of them suggested **voluntary guidelines for joint programming** (ERAC-GPC, 2010). In the guidelines a good overview was given about the state-of-art, bringing out the difficulties encountered in the **Peer Review process:**

- Peer review of proposals is at the heart of any excellence-based research policy and practice, as it forms the basis for decisions on which research(ers) will be funded. Procedures for peer review may vary across the Member and Associated States, thereby making it difficult to compare potential and achievements at the European level.
- The rationale for commonly accepted peer review procedures is most pressing in the cases when actual joint funding of research takes place through competitive calls. In those instances, commonly accepted peer review procedures are essential for a smooth management of the joint calls.
- The definition of an agreed set of evaluation criteria, among which the assessment of Excellence in Research should be regarded as the central pillar, is the basis for any scientific Peer Review system.
- It must be however recognised that divergence of approaches concerning a number of ancillary elements, including the possible use of additional non-scientific criteria, would require attention if consistency of evaluation results is to be achieved. High level of expertise among the peer reviewers is certainly a must, however quality evaluations come from diverse panels of experts, which might include a mixture of backgrounds and, if relevant, different straightforward approaches and they will usually have to be tailored to the type of call. Where necessary, experts without formal academic qualifications may be needed, for example to judge applied research with a more immediate commercial potential.
- The idea of drawing up a common database of "certified" experts needs to be treated carefully. In fact what might appear initially simple and attractive to implement, raises a number of problems (how and by whom the certification is made; how discipline boundaries are defined; how possible reputational consequences for experts who are deemed unsuitable for the database should be dealt with).
- An allied issue is that of incentives for peer reviewers. Some agencies pay their experts, while others do not. Given the limited availability of highly qualified experts, and multiple demands from different agencies, the 'market' for peer reviewers needs to be analysed, including the possible identification of non-financial incentives. g) There are usually limits to transparency: for example, while it is common practice to publish the names of the experts, this is normally done in a way that does not link individual experts to specific proposals. There may be however circumstances where the disclosure of such a link would be appropriate, as in the case of standing panels. This may also promote a sense of accountability among the experts and limit the risk that undisclosed conflicts of interest might otherwise represent.
- There are needs to be some common guidelines on what constitutes a Conflict of Interest, possibly distinguishing between what would represent 'disqualifying' and 'potential' conflict conditions, as done in the case of the rules applicable to FP7 evaluations. The cases, if any, in which Conflict of Interest conditions might be occasionally relaxed, should be also well specified.
- A suitable language regime should be established: this in most cases might boil down to the question of allowing proposals to be submitted in a language different from English. However, in case of a positive answer, further restrictions (i.e. allowing only 2 or 3 additional languages) might appear arbitrary and the practical implications of applying an open linguistic approach should be carefully considered.
- A further aspect to be considered is the way to deal with possible complaints over the peer review process, giving either no possibility of appeal, or setting up a formal redress procedure.
- While some ethical issues can be left as a matter for national regulation (for example, authorizations of clinical trials), others (e.g. use of human embryonic stem cells) are highly sensitive and potentially controversial. Agreement on the way these questions should be tackled before undertaking a common research programme.

Recommended Guidelines of the Peer Review process consisted of the list of core principles:

- Relevance – Proposals are eligible when the objectives of the specific JPI are met. The Socio-economic impact and innovation potential should be also taken duly into account.
- Excellence – The evaluation should aim at assessing the scientific excellence of the proposals. Provisions should be made towards evaluating multi-disciplinary proposals, to ensure that they are not penalised with respect to those aligned within traditional disciplinary boundaries.

- Impartiality – All proposals submitted to a call should be treated equally, i.e. evaluated impartially on their merits, irrespective of their origin or the identity of the applicants.
- Transparency – Funding decisions must be based on clearly described rules and procedures, adequately publicised. Applicants should receive a circumstantiated feedback on the outcome of the evaluation of their proposals.
- Quality – Proposal evaluation should be consistent and conform to high quality standard, similar to those achieved in other similar European or international processes.
- Confidentiality – In principle, all proposals and related data, knowledge and documents should be treated in confidence, according to established best practices.
- Ethics and Integrity – Any proposal found to contravene fundamental ethical or integrity principles may be excluded, at any stage.

Members of the evaluation panels conducting the peer review must be recognised experts, impartially chosen taking good care of avoiding any bias or conflicts of interest. Panel composition should take into account appropriate coverage of the relevant scientific and technological domains, including interdisciplinary and socio-economic aspects. It should be also, as far as possible, balanced in terms of gender, age, affiliation and nationality, including representatives from the civil society. The use of a common and certified expert database, which might be derived from the consolidation of existing ones, could be considered. All participants in a peer review panel must adhere to a Code of Conduct, which should include provisions regarding confidentiality, declaration of conflict of interest, ethical issues, as well as the sanctions to be applied in case of breach of the Code. Names of panel members having taken part in an evaluation exercise should be published after the completion of the assessment work, avoiding to associate individual names to specific proposals.
Assessment criteria should be clearly worded and defined, limited in number and logically related to the objectives of the call. The applicable marking scale, including the thresholds between fundable and non-fundable proposals, should be published with the call. Selection and funding decision should be, in principle, based on the ranking provided by the peer review experts, taking into account the budget available for each of the individual topics that might be listed in the call. Suitable **controls** should be put in place to avoid errors and ensure the fairness of the evaluation process. The outcome of such controls should be used also to improve future evaluations. It is recommended that a fast redress mechanism should be established in case of a procedural mistake occurring despite the controls put in place.

Another report, published in 2010, treated the assessment of **Europe's University-Based Research** (AUBR, 2010). It has become a major issue for a wide range of stakeholders at all levels. One of the main reasons is that research performance is widely regarded as being a major factor in economic performance. Because of their interlinked roles in education, research, and innovation, universities are considered a key to the success of the Lisbon Strategy with its move towards a global and knowledge-based economy. Improving the capacity and quality of university-based research is thought to be vitally important for innovation, including social innovation. The growing concern for the quality and assessment of university-based research partly explains the increasing importance attached to university rankings, especially global rankings.
The Expert Group on Assessment of University-based Research was established in July 2008 to identify the parameters to be observed in research assessment as well as analyse major assessment and ranking systems with a view to proposing a more valid comprehensive methodological approach. In addition different case studies were presented: a) The use of **peer review panels,** to ensure a broader understanding of the research being assessed, as well as of its contribution to knowledge, and to facilitate the assessment of research in emerging

new disciplines and of interdisciplinary research; b) The **combination** of **peer assessment** and **bibliometric indicators.**

The following general recommendations, based on good practice were made**:**

- ***Combine indicator based quantitative data with qualitative information***, for example information based on expert peer assessment. This enables the quantitative information to be tested and validated within the context and purpose of the assessment, with appropriate reference to the discipline and disciplinary practice.
- ***Recognise important differences across research disciplines***. Peer-reviewed journal articles are the primary publication channel for practically all disciplines, but the complexity of knowledge has led to a diverse range of output formats and outlets.
- ***Include assessment of impact and benefits.*** Because research does not exist in isolation, assessment should include indicators which are capable of capturing and recognising this. This differs for different disciplines. Stakeholder esteem indicators can show how research is viewed by the wider community.
- ***Integrate self-evaluation*** as a useful way to include the research community pro-actively in assessing their own contribution, but also as a means of placing the research process – which includes the organization, management, and developments over time – into context and related to institutional mission.

The Expert Group on Assessment of University-based Research has proposed a ***Multidimensional Research Assessment Matrix***. Adapting the Matrix to web-based technologies would enable different users to personalise the various dimensions and characteristics to meet their particular policy objective. This would substantially enhance its user-friendliness.

ERA-NET schemes

From the beginning of the EU 6th Framework Programme, the European Commission initiated a new project scheme, called ERA-NETs, which made it possible for research funding organisations to collaborate and to start mapping their activities. One of the first tasks was the coordination of the PRev system. Some examples from ERA-NET:

Bonus

The programme had its first call in 2007 and it was created as a result of an ERA-NET project BONUS that brought together key funders of research around the Baltic Sea to create conditions for a joint Baltic Sea research programme. One of the tasks of the project was to develop a common evaluation scheme for the future programme. This was done by organising a workshop for the partners and with the help of a questionnaire. Guidelines for a common evaluation scheme were developed as a recommendation from the partners for the future research programme. (BONUS, 2006)

The evaluation guidelines pay special attention to having common procedures for evaluating project proposals within the future joint programme. In addition, recommendations are also made concerning the mid-term evaluation and final evaluation of the programme as a whole. The programme evaluation will be carried out by an international panel that is appointed by the programme steering committee. It is recognized that carrying out the evaluation requires different kinds of expertise in addition to the scientific expertise. Thus, researchers but also policymakers, people with previous experience in programme management and other stakeholders are needed. It is stated in the guidelines that clear and measurable goals should be set in the planning phase of the programme, and that the goals should be measurable during the programme, at the end of it and some time after the programme have ended. According to the guidelines, it is important that the various partners are unanimous about the goals and how to measure them. (BONUS, 2006)

Social sciences and humanities: HERA and NORFACE

Good examples come from the area of social sciences and humanities where one of the first tasks was to map the best practices in peer review mechanisms across HERA and NORFACE members (HERA, 2005; NORFACE, 2005, 2006). In the study about the Evaluation and Benchmarking of Humanities Research in Europe, the authors (Dolan, 2007) spotlighted **six reasons** why pooling resources in peer review may be beneficial:

1) it can be argued that excellent peer review processes often require very detailed and/or location-specific expertise that is not readily available in specific countries. Furthermore, it can also be argued that very specific expertise may involve national communities that are relatively small, in which it becomes difficult to avoid conflicts of interest;

2) it seems unusual that in a few decades of European research collaboration there is still nothing resembling a common standing European database of peer reviewers, which includes data commonly used by all relevant funding agencies. Yet, expertise is like many other things: a rare resource which ought to be pooled so as to maximise its impact and enhance its sustainability;

3) some funding agencies across Europe seem to be having difficulties in obtaining peer review reports due to 'peer review fatigue'.

4) Suggestion was made to use the Peer Review College model which could help to disseminate best practice in peer review across Europe, and contribute towards the creation of a specific and common understanding of notions such as 'quality', 'excellence', 'interdisciplinary', etc.

5) it was stressed that the more European funding organisations have a common European peer review 'culture', the easier it will be to make use of each other's expertise. This would allow all relevant agencies to efficiently pool resources and use each other's peer reviewers.

6) the agencies could use the internationalisation of peer review as a metric or indicator of quality – the more internationalised the process becomes; the more agencies can state that it is operating on the basis of international standards or 'world class' peer review.

In order to facilitate the exchange of good practices and make available the wealth of experience matured within the ERA-NET scheme, the European Commission (EC) set up the ERA-NET Learning Platform (a support action started in 2009), which will produce a call implementation toolbox and a set of recommendations for evaluation standards and funding modes. The EC and, more recently, the European Research Council (ERC) have developed also a lot of direct expertise in organising peer reviews in the context of implementing the successive Framework Programmes.

European Research Council (ERC)

The main goal of the European Research Council (ERC) is to encourage high quality research in Europe through competitive funding. The selection of scientific and scholarly proposals for ERC funding is based on international peer review with excellence as the sole criterion. The ERC uses a typical panel-based system, in which panels of high-level scientists and/or scholars make recommendations for funding. The panels of each grant are grouped into three disciplinary domains that cover the entire spectrum of science, engineering and scholarship: a) Social sciences and Humanities (SH); b) Life sciences (LS); c) Physical and Engineering Sciences (PE). Research proposals of a multi- and inter-disciplinary nature are strongly encouraged throughout the ERC's schemes. Proposals of this type are evaluated by the ERC's regular panels with the appropriate external expertise. Each ERC panel consists of a chairman and 10–15 members. The Panel Chair and the Panel Members are selected on the basis of their scientific reputation. In addition to the Panel Members (who act as “generalists”), the ERC evaluations rely on input from remote experts external to the panel, called referees. They are scientists and scholars who bring in the necessary specialised expertise. Before the deadline of a call, the

names of the panel chairs are published on the ERC website. Similarly, the names of panel members are published, however, after the evaluation process is concluded. **(ERC)**

European Science Foundation initiatives

ESF and EuroHORCs have been studying the peer review issue since 2006 and included it in their strategy document "**Vision on a Globally Competitive ERA and Road Map for Actions**" (Vision, 2009) where the two organisations propose to establish European-level benchmarks for peer review processes, to set up European peer review panels and to develop European-level peer review resources, such as quality-controlled shared databases of reviewers.

In 2006, ESF organised a conference in Prague (ESF, 2006) with the aim of analysing contemporary trends in the evaluation of research, examining how the peer review process is understood and performed, and considering its future modifications. As a general focus, the theme of the conference was organised around three questions:

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

In 2007 the German Research Foundation (DFG) proposed to the European Science Foundation to establish a Member Organisation Forum on Ex-Post Evaluation of Funding Schemes and Research Programmes (MO Forum) with the **objectives:**

- To facilitate networking of science officers engaged in evaluation;
- To exchange and document experiences with current practices;
- To explore needs and possibilities for collaboration in future evaluation exercises.

From October 2007 to April 2009, the ESF Member Organisation Forum convened four workshops on the following focused topics:

- Evaluation of Funding Schemes and Research Programmes: Expectations, Practices and Experiences;
- Quantitative Indicators in ex-post Evaluation of Funding Schemes and Research Programmes;
- Best Practices in “Quality Assurance”;
- Socio-economic Impact Assessment.

On 1 October 2009, the Final Report of the Forum was delivered (MO Forum, 2009). One of the goals of the Forum was to develop a mapping of current evaluation activities within research organisations. **Five types of evaluation** were identified as most common in most organisations:

1) Evaluation of the funding agency as an organisation;

2) Evaluation of funding policies (or particular strategic issues);

3) Evaluation of research fields or scientific disciplines;

4) Evaluation of funding schemes;

5) Evaluation of research grants (to a single Principal Investigator or a group of recipients).

In addition, case studies of selected funding agencies were presented (German Research Foundation – DFG; Research Council of Norway – RCN; Austrian Science Fund – FWF; Netherlands Organisation for Scientific Research - NWO; Swedish Research Council – SRC). In conclusion, some observations were made:

- Different terminology relating to funding activities and evaluation exists among agencies which hampers comparison and understanding.
- Notwithstanding the different terminology, funding schemes are generally comparable across funding modes.
- Most scheme evaluations employ a mix of quantitative and qualitative methodologies and are typically conducted by dedicated evaluation practitioners.
- Differences in the size, age and structure of organisations have an impact on the evaluation practices employed.
- It was not uncommon for agencies to develop schemes without giving careful consideration as to how success and impact of the scheme would be measured in the future.

Evaluation of individual grants
Ex-post evaluation of individual research grants through the use of final grant reports was a common strategy in research funding agencies. Analysis of practices by the Forum's participating organisations showed that final reports are collected by almost all funding agencies and are considered an integral part of the funding process, marking the "judicial" end of the funded grant. Final reports are also increasingly used as a source of information concerning the results of funded research beyond the single project, to the level of funding schemes and funding strategies.

In March 2011, MO Forum published the **Peer Review Guide** (MO Forum, 2011) which describes the practices across the members of ESF and EUROHORCs by setting a minimum core of basic principles on peer review processes commonly accepted at a European level. The content of the Guide is structured according to **three thematic areas**:
1) an introduction to peer review in a general sense, and a typology of funding instruments, also the pillars of good practice in peer review;
2) peer review methodology;
3) description of the variants of the funding instruments and their implication for peer review.

Characterising the appropriateness of peer review practices can be meaningful only when considered in the context of the specific programmes or funding instruments to which they must apply. Therefore, in order to establish common approaches and understanding of the practices of peer review, it is necessary to establish common definitions and meanings in the context in which they are to be used. This context is defined by various funding opportunities with specific objectives that different organisations have developed in order to select competing proposals and to allocate merit-based funding using clearly defined objectives and selection criteria. Peer Review Guide was based on a survey in which thirty organisations from 23 European countries, one from the United States of America and several supranational European organisations participated, the results of the survey were published (MO Forum Appendix 2, 2011).
All main research fields were covered by at least 80% of the organisations. The two funding instruments mostly utilised by the majority of the organisations were *Individual Research Programmes* (90%) as well as *Career*

Development programmes (87%). The majority of the organisations received up to 5,000 proposals per year. From all responding organisations, 36% reported an average success rate of 21–30%.
On May 14–15, 2012, the Global Summit which brought together the Heads of Research Councils from research intensive countries was held in Washington DC. The Global Research Council has a long-term objective of fostering multilateral research collaboration across continents to benefit both developing and developed nations (http://www.globalresearchcouncil.org/). The goal for the Washington Global Summit on Merit Review was to develop and endorse a Statement of Principles on Merit Review to identify best practices and standards that will cultivate multinational research cooperation among countries and across continents**. Six main principles** obligatory in Peer Review Process were agreed upon:
1) **Expert Assessment** – Collectively, reviewers should have the appropriate knowledge and expertise to assess the proposal both at the level of the broad context of the research field(s) to which it contributes and with respect to the specific objectives and methodology. Reviewers should be selected according to clear criteria;
2) **Transparency** – Decisions must be based on clearly described rules, procedures and evaluation criteria that are published *a priori*. Applicants should receive appropriate feedback on the evaluation of their proposal;
3) **Impartiality** – Proposals must be assessed fairly and on their merit. Conflicts of interest must be declared and managed according to defined, published processes;
4) **Appropriateness** – The review process should be consistent with the nature of the call, with the research area addressed, and in proportion to the investment and complexity of the work;
5) **Confidentiality –** All proposals, including related data, intellectual property and other documents, must be treated in confidence by reviewers and organizations involved in the review process;
6) **Integrity and Ethical Considerations** – Ethics and integrity are paramount to the review process.

National Peer Review studies

The work of European Science Foundation is mostly based on the experience of their member organisations. In this respect, in this section we point out a few studies conducted on the national level. New tendency, characteristic of the 21st century, is involvement of the interests of society as a new component into the Peer Review agenda. In 2004, the outcomes of the special survey on equipping the public with an understanding of peer review were introduced in the UK (Brown, 2004). The Working Party recognised that scientific peer review has not traditionally been a subject of public interest. In the British society, however, science has become the subject of many wider public and political controversies. More scientific information is being put into the public domain and a growing number of organisations are becoming involved in promoting and discussing scientific research and reacting to new research claims. Scientific evidence is sometimes mixed up in these 'politics of science'. Exaggeration and anxieties about scientific developments often relate to research findings that are regarded by scientific experts as weak or flawed, or that have not been subjected to independent expert scrutiny at all. These developments have resulted in a greater public need for clarity about the status of new research claims. A wider understanding of peer review's role, in assessing whether work is competent, significant and original, is central to achieving that clarity about the status of research. The opportunity to explain peer review needs to be seen within this *broader social interest* in the reliability and quality of research, rather than identified with the preoccupations of particular scientific groups that their messages are not getting through.
The public, in its widest sense, should be encouraged to ask questions about peer review when listening to claims about scientific findings in an interview, press release, or news report. Has the work been evaluated by experts in the field, or is the report based on opinion or unsubstantiated extrapolation? Is it acknowledged by other scientists as a contribution to the field, or dismissed because it is flawed? Has it been replicated?

One of the reasons for establishing a Working Party was the predominance of criticism of peer review, relative to the paucity of explanations about what it is or why it has become the system for sharing scientific findings between scientists. These criticisms are often concerned with very different things. Scientists tend to be concerned about the practical difficulties involved with managing the peer review of thousands of papers and maintaining standards. Some individuals are called upon very frequently to review papers and attentive reviewing takes time. Some critics of scientific practice have accused peer review of being a barrier to new ideas. The Working Party has looked for ways to promote a culture in which people who promote research claims in the public domain feel obliged to explain the standard of the evidence on which they are based. In doing this they can encourage the public to ask more effective questions about the scientific information put before them.
Some recommendations on how scientific information enters the public domain and interact with the peer-review process: a) Scientists' conferences, press releases and other promotional activities should help this process by stating clearly whether particular scientific claims have been peer reviewed wherever possible; b) Scientists should work with press officers to ensure that their peer reviewed work is reflected accurately in all publicity. Universities and other organisations involved in scientific research should ensure that any press officers who do not have a background in science understand the peer-review process, for example as described in this discussion paper; c) It is further recommended that scientists follow reports of their research in the wider media, and try to correct claims that deviate substantially from peer-reviewed results; but that they distinguish between this and matters of taste and style in how others choose to discuss their work; d) it is recommended that conference organisers try to put information about the peer-review status of claims into their promotional literature, and encourage presenters to communicate with them about this when (i) a talk is clearly likely to cause wider controversy; or (ii) new findings are being widely promoted to draw attention to a conference; e) A best practice guide should therefore be developed by companies that are obliged immediately to report R&D results to the financial markets and to product licensing authorities. It is also recommended that the use of an 'open access' Web-based resource be explored, where organisations can provide supporting scientific data simultaneously with any press release; f) It is recommended that bodies concerned with devising curricula, producing teaching materials and promoting science education, produce teaching resources on peer review for educators for all age groups.

The Working Group of British Academy on Humanities and Social Sciences (British Academy, 2007) was established to examine how the practice of peer review functioned in a context in which its scope was expanding beyond its traditional primary focus on individual publications and grants to encompass broader evaluations of, say, the research performance of departments. These kinds of evaluations, of which the most important in the UK has been the Research Assessment Exercise (RAE), typically involve assessors examining in a summary way publications and grant-aided research that have already been peer reviewed, so they can be thought of as second-order or secondary peer review. Particular attention was paid to issues specific to the humanities and social sciences. In the report following, questions were tackled as the way in which peer review is conducted and will be conducted in the future. Some of these developments are positive (such as advances in information technology, which speed up the process and also make it easier to use international referees). Others pose particular challenges (such as the increase in the volume of submissions both for grants and for journals), which add to the burdens of peer review. The proposed development of metrics to play a more prominent role in the assessment of the research performance of departments (RAE) interacts in complex ways with the practice of peer review. Peer review in practice takes a wide variety of forms, reflecting the diversity of subject matter and approaches in humanities and social science research. There is no one model that all should follow – the peer review is not in fact a single process, but rather a flexible set of mechanisms. This variety of practice is important

in relation to publication. There are many different models of peer review used. It is a considerable merit of the way in which the peer review works in journal publications that there is not one single model of good practice that all should follow, but instead decentralised diversity. Nevertheless, there are principles that good peer review should follow. These include **timeliness, transparency** and **verifiability**. These principles cannot guarantee the identification of the best quality work on a fair basis, but without them quality and fairness will suffer.
In the case of grants peer review remains essential if good work is to be identified. In a situation in which applicants have few alternatives to funding, it is important that funding bodies uphold the integrity of their peer review processes. It is also important that they find ways of responding to the innovative and the risky. The following **recommendations** were made:

- **Training** – training guidelines should be amended so that explicit attention is given to the need to train postgraduates in the norms and conventions of peer review. Training should also be made available to early career track postdoctoral researchers;
- **The Costs of Peer Review** – Surveys and other evidence have shown that there are various reasons why academics participate in peer review. Not all motivations are altruistic, and there is no reason why they should be. However, a central element, without which the peer review system would not exist, is the professional commitment to contribute to the academic public good. Each university in receipt of public funds should accept an obligation to encourage its researchers to engage in these activities, recognising that peer review is an essential part of the fabric of academic life – the costs of which are met by the funds allocated by the Funding Councils to support research infrastructure. Develop a more sophisticated understanding of the costs of peer review, and maintain efforts to ensure that the peer review processes are proportionate to the scale of award;
- **Metrics and peer review** – Metrics and peer review are related in complex ways. Metrics are typically summary measures of peer reviewed activities, and the use of metrics sets up incentives to change behaviour in ways that affect peer review. Care should be taken to ensure that any metrics employed reflect the distinctive nature of the humanities and social sciences research and do not have an adverse effect on the quality of the work that they are seeking to measure;
- **Peer review and innovation** – the variety of practices in the conduct of peer review in the humanities and social sciences is a strength, not a weakness. Research funders should take pains to avoid a mechanistic approach in their decision-making processes for the award of research grants in the humanities and social sciences. Consider grant proposals on a case by case basis, taking pains to ensure that award decisions are not made by the application of formulae.

Issues of knowledge transfer and impact play an important role in public policy, and are likely to become more, not less, important over time. Quality should not be sacrificed in favour of relevance and impact. **Applied research ought to meet the same standards of research design, sample selection and evidential inference that applies to any sort of work** (allowing for the practical difficulties of conducting applied research). Indeed, if research is being used by policy makers to take decisions on matters that have a direct effect on the quality of citizens' lives, **the standards ought to be as high as possible**. Similarly, **novelty cannot be regarded as a substitute for quality**. Ensure that considerations of applicability and relevance do not compromise judgements of quality. Set aside funds for risky, speculative projects. Ensure that there is a healthy **balance between strategic funding and responsive mode projects**. Encourage endowments within universities to support small grants for innovative, high risk research. Ensure that the **process of selecting topics for strategic initiatives is also subject to peer review.**

A good example of how to develop the PRev system comes from the **Netherlands**. The Standard Evaluation Protocol 2009–2015 (SEP) is the fourth protocol for evaluation of scientific research in the Netherlands, following the protocols of 1994, 1998 and 2003 (SEP, 2003; SEP, 2009). The aim of the SEP is to provide common guidelines for the evaluation and improvement of research and research policy, based on expert assessments.

This process is under permanent assessment, and for example in last evaluation of SEP the suggestion was made that more emphasis should be placed on societal relevance, on positioning and on benchmarking.
The external evaluation of scientific research applies at two levels: the research institute as a whole and its research programmes. Three main tasks of the research institute and its research programmes are to be assessed: the production of results relevant to the scientific community, the production of results relevant to society, and the training of PhD students. Four main criteria are considered in the assessment: **quality, productivity, societal relevance** and **vitality**, and **feasibility.**
Praiseworthy work has been done by Hanne Foss Hansen (Foss Hansen, 2009) giving the insight into methods for research evaluation, charts the practice of research evaluation and developmental tendencies and challenges in this area. Various types of peer review are presented – classic peer review, modified peer review and extended peer review – and their strengths and weaknesses are discussed. The capacity for various research indicators to provide evidence is elucidated in so far as particular emphasis is placed on tracking numbers of publications and citations. Methods are presented for both ex-ante and ex-post assessment of the relevance of research and its societal quality. Examples are also given of more integrated models for research evaluation. An illustrative survey of the practice of research evaluation internationally and nationally is provided. The survey places focus on evaluation at four levels: the level of systems, the level of programmes, the level of institutions and operators and the level of institutes and groups. Concrete examples are given of evaluation and evaluation systems from a range of countries. The survey also includes a presentation of evaluation practice in the EU and OECD, and the ever more widespread use of ranking lists at the university level is discussed. It is shown that new methods are still being developed – more variants of peer review, new types of research indicators, methods for the assessment of relevance and societal quality and increasing numbers of integrated models. The analysis also shows that no ideal method can be indicated for research evaluation. **All methods have their potential and their limitations.** It is often fruitful to use a **mixed-method approach**.
The Nordic countries have a long-standing tradition of using peer review for the evaluation of proposals for research funding, as well as of retrospective evaluation of different research programmes and disciplines or fields of research (NordForsk NORIA-net, 2010). NORIA-net project aimed at a) identifying current practices of peer review in research evaluation and research applications; b) identifying the main challenges and biases related to the use of peer review methods; c) finding new ways to minimise the challenges as regards peer review methods; d) improving the quality and efficiency of evaluation procedures by identifying and implementing good practices; e) addressing the challenges of using international reviewers by planning joint Nordic panels; f) becoming more competent and efficient in managing research funding by advancing collaboration between Nordic partners and by joining forces to evaluate research proposals, and g) organising a joint peer review exercise. The main recommendation was to discuss the possibility to establish a Nordic database of experts.
The comprehensive overviews about PRev in Health Sciences have been made in the **RAND Reports series** (Ismail, et al, 2009; Wu, et al, 2011). PRev has been taken as premier means for assessing the quality of research proposals, but there are limitations, which must be considered. PRev cannot be a panacea, and there may be better ways of allocating research funding if the aim is to fund highly innovative work, to support early-career researchers, or interdisciplinary research. There is an urgent need for better understanding of the reliability and fairness of peer review in particular, even if conclusive evidence on its ability to fund the 'best' research is unlikely ever to be forthcoming.
Potential modifications to the grant peer review process may be considered to improve efficiency or effectiveness. With respect to efficiency, for example, improvements could be brought about by *moderating demand* to ensure that the number of applications received is kept below a certain threshold – thus reducing the burden on reviewers and applicants. This could be achieved by (i) reducing advertising; (ii) changing deadline systems for

funders that use fixed milestones for submission; or (iii) limiting the number of applications from particular institutions. It may also be possible to *streamline assessment procedures* using tighter systems of triage on applications received. Other potential cost-saving measures include (1) reducing the number of external referees involved in peer review of grant applications, and (2) increasing the use of technology – including videoconferencing – so that peer review panellists do not have to gather in one place for scoring meetings. Different kinds of peers should be used for different purposes – specifically targeting specialists in translational or high-risk, innovative research, for example, where this is the desired outcome. This has important implications for funding bodies; since reviewers both identify and define good research, an extensive understanding of different views within a field will be required by the person selecting reviewers.
Improving the capacity of peer review to support applied research**:** Panel members are drawn from both academic peer review and decision making constituencies; but educators and communication experts may also participate if the proposal in question is likely to be a high-impact area of research. The aim is thus to evaluate research proposals both in terms of their scientific merit and the potential impact they may have. Improving the capacity of peer review to support innovative research, s.c DARPA model: a narrowed down version of peer review, in which there is no panel, simply 'expert' judgement by a specially selected programme manager.

Pros and cons of Peer Review Process versus Metrics

The basic argument for bibliometrics in grant reviews is that research councils need to evaluate whether or not public funds have been well spent (Oswald, 2010) and the prioritization of the need to evaluate economic and societal relevance of research (Hicks, et al 2004; Scott, 2007). The argument essentially debates the viability of use of quantitative data (metrics) versus qualitative social judgments (peer review) in measuring research quality (David, 2008).

The key metric indicators are Citation Indices and Impact Factors of journals (Lundberg, 2007), articles and research outcomes. Indicators based on bibliometric methods offer much more than 'only numbers'. They provide insight into the position of actors at the research front in terms of influence and specializations, as well as into patterns of scientific communication and processes of knowledge dissemination (van Raan, 1996). Bibliometrics is seen as providing a reliable, objective and cost effective contrast to the subjectivity of peer review (Glaser, 2004), is neutral and allows comparative (national and international) assessment (Oxman and Guyatt, 1991) and has in the past few years been used in Australia and the UK in the evaluation of the research of individuals (Levitt and Thelwall, 2011). The UK's Research Assessment Evaluation exercise has since 2008 used indicators based largely on publication and citation data (Adams, 2009), however much of the data do not reflect on the quality of the research. Consequently peer assessment should not be wholly replaced by bibliometrics (Richards, et al, 2009) but should be complementary. Research indicates a correlation between peer review and the most popular bibliometrics (h-, m- and g- indices, total citations and mean number of citations) but does not explain the high number (40%) of variances (Lovegrove, and Johnson, 2008). The suggestion is that a synergy between peer-review and the various types of bibliometric indicators is necessary (Lovegrove, and Johnson, 2008.; Lundberg, 2007).

Research favouring the use of bibliometrics suggests that it is highly correlated with various quantitative indicators and should be used to compile a journal quality index (Taylor, 2011), has potential as a tool for

evaluation (Haeffner-Cavaillon and Graillot-Gak, 2009) but should be used in conjunction with the PRP (van Raan, 2000). The h-index is the most widely used criterion (Bornmann and Daniel, 2007) but while correlating closely to peer judgments is more effective evaluating basic science applications and less effective in specific (Rinia, et al, 1998) and small research fields that have a smaller number of citations (van Raan, 2006). By contrast, whereas PRP is less effective in small fields, (Abramo and D'Angelo, 2009; So, 1998) concluded that bibliometrics works better in small research fields and for better known scholars. Bibliometrics has a particular use in comparative evaluations of senior faculty research performances (Meho and Sonnenwald, 2000) as well as for assessing interdisciplinary research for which (Rinia, et al, 2001) found neither any bibliometric bias nor it has to be said peer-review bias.

Research critical of bibliometrics base their arguments on the proverb of 'change for the sake of change' which suggests that a system that has lasted for three and half centuries will develop flaws but should be developed as opposed to being totally replaced by a system that (Molinie and Bodenhausen, 2010) refers to as 'the tyranny of bibliometrics'. Indeed (Bloch and Walter, 2001) argues for the abandonment of all Impact Factor related indicators and a return to the basics of peer-review judgments. One element of the tyranny is the reliance of citation-based indices (e.g. the h index) on research output being published in the ISI/Web of Knowledge journals of Thomson Reuters (Garcia-Aracil, et al, 2006; Franks, et al, 2006) and Impact Factor data produced by Thomson Reuters (Pendlebury, 2009). An additional critique of citations comes from a Norwegian study (Aksnes, et al, 2011) that indicates that female researchers are cited considerably less frequently. Particularly as some highly cited papers do not accumulate a large number of citations in the initial three years after publication on which many citation indices are based (Allen, et al, 2009). Research output that is published elsewhere is automatically disregarded as being 'not excellent'. Norris and Oppenheim (Norris and Oppenheim, 2010) suggest that detailed cited reference searches must be undertaken to rectify this failing. Scientists are not alone in viewing bibliometrics in a critical light. Accounting scholars in the UK place more value on peer reviews than on bibliometrics (Brinn, et al, 2000). As already stated by Foss Hansen (Foss Hansen, 2009), no ideal method can be indicated for research evaluation. All methods have their potential and their limitations. It is often fruitful to use a mixed-method approach.

Peer Review practices

As we see from papers of Georgiou Georghiou, L. (1995) and Luukkonen (Luukkonen, 2002), the prerequisite of good peer review is its permanently improving character, and involvement of high level stakeholders and experts. The fact that there is no single European or global way of doing research evaluation is seen as a problem by several researchers. As yet no ISO applies to any part of the peer-review process as regards either the publication of research output or the granting of research funding. This is essentially the point, Langfeldt (Langfeldt, 2001) makes concerning the non-existence of any norms for assessments and particularly which criteria reviewers should focus on. Consequently the outcome of a grant application may be 'accidental'. The

relevance of ISOs applies particularly to Reviewers, who while forming the largest unremunerated element of either arena, are highly influential in the quality of publishing output as well as the successful outcome of grant applications. Traditionally reviewers have been 'part-and-parcel' of the system applying their expertise in the reviewing process without experiencing any training. The overall reliability of reviewer judgments is recognized as being extremely poor (Jayasinghe et al, 2006), although Marsh et al (Marsh et al, 2008) discovered the only major systematic bias occurred when the applicants could nominate their assessors.

So which indicators of research articles and grant applications are important to reviewers? Bazeley (Bazeley, 1998) in an Australian context of grant funding examined the effect of both societal variables as well as institutional variables plus the history of applicants' success or failure in the grant process. The strength of the proposal was deemed to be the most significant variable, the others having only limited or partial significance. The most interesting conclusion was that the 'Matthew effect' of accumulative advantage (gained through review and assessment of each variable) was quite significant. On the other hand, Canibano et al (Canibano et al, 2009) discovered that the key (indeed only) influence was the research productivity of applicants.

In Canada, in addition to the National Research Council there are federal agencies that grant funding to federal universities, known as 'granting councils' (Holbrook, 2000), which use the unusual method of reviewing the success of the principal investigator in carrying out a previous project.

The number of reviewers is critical. Although the Working Group on Peer Review of the Advisory Committee to the Director of NIH has recommended that at least four reviewers should be used to assess each grant application, Kaplan et al (Kaplan et al, 2008) found that a significantly larger number is necessary to meet the precise criteria set by the NIH. A study (Herzog et al, 2005) that reviewed different types of abstracts (for research presentations, for applied human-animal programs and as critical reviews), rated the reliability of judgments of teams of three reviewers. The reliability was poor for the first two categories and fair for the last. The best results were achieved by averaging the ratings of the three reviewers and using their combined scores to make judgments.

There should not be any difference between internal reviewers and external reviewers, yet Hodgson (Hodgson, 1995) discovered that internal reviewers more closely matched a review committee's final decision than external reviewers.

A major problem in peer-review judgments is the substantial error that occurs when a large number of reviewers evaluate only a few articles or grant applications. To overcome this problem, Jayasinghe et al (Jayasinghe et al, 2006) trialled a 'reader system' in which a select few senior academics read all the articles or grant applications in their particular field of scholarship. This system produced greater reliability of judgments for both the research and the researcher and meets the approval of the study conducted by Marsh et al (Marsh et al, 2008).

The team of LiquidPub project (Birukou, et al, 2011) presents their ideas of several novel approaches for research evaluation (bidding as an indicator of importance, peerevaluation.org vs. ucount, use of community opinions, incentives to participate, the role of the internet) which are alternative or complementary to traditional peer review. One of the conclusions that authors draw is that, as the landscape of the scientific publishing is undoubtedly changing, the processes for the evaluation of research outputs and of researchers are also changing. In the next years we envision the growth of various tools for research evaluation, including open source and those operating with open API/protocols. Such tools would primarily operate on the Web and include the variety of methods for research evaluation, so that PC chairs or journal editors (or even people playing some new emerging roles which do not exist yet) will be able to choose. Examples of tools with such functionalities already emerge (e.g., Mendeley, Peerevaluation.org, Interdisciplines) but it is not yet clear how these tools can be connected and which of them will be adopted widely enough to have a normative effect. Attention should be paid less to designing "the" scientific evaluation system of tomorrow – something that, like "the" peer review process,

will be an emergent phenomenon based on the different needs of different disciplines and communities. Instead, attention should focus on ensuring interoperability and diversity among the many possible tools that scientific evaluation can make use of.

Peer Review biases

Bias issues, in the contexts of the peer review process (PRP) of research articles and in the assessment of research grant applications, share both a similar time-frame and topics of concern. Bias was in all likelihood not an issue of concern in the initial two hundred and eighty years following Henry Oldenburg's adoption of the peer review process at the Royal Society of London in the 17th century (House of Commons Select Committee Report – HCSCR). During this period, 1665–1945, publishers of academic research are unlikely to have felt concern about either societal or institutional issues of bias if only because published academic research was dominated by male researchers working at the worlds' foremost universities that were mostly to be found in the West (Europe and North America).

The societal changes engendered by the Second World War and the global expansion of tertiary education research institutions broadened the gender and ethnographic base of researchers. Consequently, bias became a 'hot' topic from the 1950s onwards in the PRP of research articles.

The same time-frame is equally valid for the use of PRP in assessing research grant applications as it was not until the inter-war years (1918-1939) did funding institutions adopt PRP (in the USA).

The editor of The Lancet shares the view with (Atkinson, 2001) that the academic research community knows that the peer review process (PRP) is biased. Just as proving the existence of bias is not easy and does not necessarily provide a homogenous result, providing a method of correction is difficult. The problems are partly a matter of personal perception of bias, individuals' ethics and personal morals; partly whether or not personal 'anecdotal' evidence can claim any validity and also a paucity of detailed research institutions' records. Surveys and questionnaires have to rely on the concept that the respondent tells the truth, and participants in PRP (authors, reviewers, editors of journals, grant assessors) are unlikely to admit to having any bias towards their peers. Research institutions do not, as a rule, keep a detailed record of their researchers' attempts (success or failures) at publishing their research or gaining research funding. Research into PRP bias is therefore dependent on analysis of the data of societal indices (gender, age, creed, and ethnicity) and institutional indices (institution, discipline, research status) held by either the research journals or the research funding institutions.

Institutional practitioners of PRP, the journals and research councils, have developed two 'anonymous peer review' approaches to counter charges of bias – the single blind review and the double-blind review.

The blind review: in which the institution (journal or research council) keeps the identities of the reviewer/s secret but does not require the author/s to remove references to themselves from the text (or application).

The double-blind review: in which the institution (journal or research council) not only keeps the identities of the reviewer/s secret but also requires the author/s to remove all references to themselves.

Neither approach is abuse proof because both depend on the assumption that that the reviewer will withdraw from the process in the event of: knowing the author/s, having a conflict of interest (i.e. a parallel path of research), having any erstwhile prejudices towards the author/s. Failures by reviewers to withdraw due to any of these are difficult to substantiate.

The alternative approach to counteract the inherent 'bias' of the traditional anonymous peer review process is the open review approach. This version not only does not allow anonymity to either authors or reviewer/assessors but also, if conducted in conjunction with the Internet, enables real-time comments and feed-back.

Gender bias

The most perceived aspect of bias, in both the journal and grant application contexts, concerns gender (women), for whom success is frequently linked to cooperation with male senior researchers (Olsson, 1999). A study to detect any gender bias at the Journal of Epidemiology and Community Health examined the under use of female reviewers and concluded that in light of an empirical evidence of gender bias there is not any justifiable argument in quality criteria for not promoting gender parity in reviewers. (Davo, et al, 2003; Scantlebury, 2002) concurs with this finding in the context of the lack of female high level reviewers in the field of chemistry (in the USA). Grant funding institutions are not immune to gender bias with two studies finding that Reviewers on the Swedish Medical Research Council were either gender biased (Wenneras, 1997) or promoted nepotism (Sandstrom and Hallsten, 2008). Indeed the issue of gender bias is sufficiently topical that Bormann et al (Bormann et al, 2008) have developed a number of approaches using a generalized latent variable modelling approach and the latent Markov models (Bormann et al, 2009) to detect bias in both the journal and grant application contexts, often focusing on gender. A generalized study of bias in doctoral and post-doctoral research applications discovered that there were statistical differences for gender, major field of study and institutional affiliation (Bornmann and Daniel, 2005), but not for nationality. While Bornmann et al (Bornmann et al, 2007) in a meta-analysis of 21 studies of research applications, conclude that male researchers have a statistically significant better chance than their female peers of success, Marsha et al (Marsha et al, 2011) in a large study of 10,023 reviews by 6233 external assessors of 2331 research proposals argue support for the null hypothesis of any gender bias. Similar conclusions were made in the assessment of the grant peer review system in Spain (Canibano et al, 2009). The survey revealed that the selection process mainly relies on the research productivity of applicants and not on other variables not directly related with scientific excellence such as gender, age or residence. A Norwegian study (Aksnes et al, 2011) shows that that publications of female researchers are less cited than those of men, although the differences are not large. The gender differences in citation rates can be attributed to differences in productivity. There is a cumulative advantage effect of increasing publication output on citation rates. Since women in the study published significantly fewer publications than do men, they benefit less from this effect. The study also provided results on how publication and citation rates vary according to scientific position, age, and discipline. Australian Research Council Large Grants Scheme survey (Bazeley, 1998) aimed to determine the extent of influence of biographical and academic “track record” variables on ratings by independent assessors and on final outcomes as determined by the Council’s discipline review panels. The variables considered included age, gender, type and status of position, institutional base, previous grants history and publication records of the applicants. Age and gender of the applicants did not have a significant impact on outcomes. Those in research only positions were rated more highly, and those in more prestigious institutions were more likely to win support, but these factors were partly accounted for by the more impressive publication records of these groups. The academic status of the applicant was, however, found to make a significant contribution to an explanation of the variance in independent assessor's ratings of the strength of the researcher, along with but additionally to that contributed by publication components of their track record. These results lend some support to the idea that the "Matthew effect", or theory of accumulative advantage, has some impact on peer and panel review of applications within the highly competitive arena of Australian Research Council large grants.

Matthew effect

"Matthew effect" was a term proposed by Robert K. Merton (Merton, 1968; Merton, 1988) to describe how, among other things, eminent scientists will often get more credit than a comparatively unknown researcher, even if their work is similar; it also means that credit will usually be given to researchers who are already famous. It seems that this is becoming a key issue for the peer review. Starting from its mentioning by Merton the number of

publications about "Matthew effect" has doubled every ten years in the 20th century, and quadrupled in the 21st century (WoS).

Field of study

Field of study seems particularly relevant in the instance of a discipline having a non-scientific branch (clinical medicine vs. homeopathic medicine) or having an innovative branch (innovation journalism in media and communication studies and organic farming in agricultural sciences). Scientific Reviewers of grant applications for organic farming was the subject of a study in Sweden (Rasmussen et al, 2006) which found that scientists experienced in organic farming provided contrasting reviews compared to their peers who did not have that experience. In a field related study (Blackburn and Hakel, 2006) the correlation between experience (authorship / non-authorship of conference posters) and degree of criticism was tested with authorship surprisingly generating higher levels of criticism.

Institutional affiliation

Institutional affiliation might seem crucial – old ('ivy league') *versus* new ('red-brick'), state funded (national research councils) *versus* private funded (Max Planck Institutes), EU15 (UK, Germany, France) *versus* EU25 (Estonia, Latvia, Lithuania) – in both the publishing arena and the grant application. While (Sofronis et al, 2011) study of Research Assessment Evaluations in the UK of Economic departments did not find any bias towards 'old' universities compared to the 'new' ones, (Oswald, 2010) did find that less well known departments in the UK produced most of the high-cited research articles and (Hegde, 2009) in a study of US National Health Institutes did find an element of 'political bias in funding' between institutions in states that had members sitting on the House Appropriations Committee and those that did not have any direct representation. The situation in a "scientifically small" country may be different. A comparative analysis of blind and open review in Korea (Lee et al, 2002) showed that the applicant characteristics (rank of undergraduate school where the applicant studied, professional age of the applicant, and academic recognition of the applicant) are the major factors leading to the significantly different evaluation scores between blinded and sighted reviewers. This means that "open" evaluation of research proposals is obviously biased.

Language

A far less obvious bias that has yet to be researched concerns the issue of writing in the language of science (English). This applies to both writing research articles as well as writing transnational research proposals. The proposed hypothesis would be that non-native English speakers and institutions that cannot afford costly language translating and language editing services are on the receiving end of a skewed system. For all that the argument is highly contentious, Bordage (Bordage, 2001) did discover in a research of reviewers' comments that a key reason reviewers gave for accepting a research manuscript was 'excellence of writing' whereas a 'poorly written manuscript' was a key reason for rejection.
Within the context of research publication and dissemination (i.e. the journals) there does not seem to be a workable solution other than a PRP based system. The point being that journals cannot use bibliometrics to determine who (which research) gets published because there is not any data for novice (unpublished) researchers.

Transparency

One viable alternative is the 'open review' developed on-line by the British Medical Journal (van Rooyen 1999, 2001, 2010). The prevalent view is that the Internet can be used to improve the transparency and quality of the peer-review process (Fitzpatrick, 2010). There is also hope that the emergence of the open peer-review process may foster a culture of "exchange of acceptance," whereby the author only selects favourable reviewers in (Slesser and Qureshi, 2009).

Time

The master thesis made at the University of Tromsø (UiT) (Refsdal, 2010) focuses on how much time is spent on reviewing at this particular institution, in addition to examining the researchers' opinions on several aspects of peer review. The method used for data collection was a survey, with an online questionnaire sent to researchers at UiT. Key findings from the survey are that 69% of the respondents are active reviewers (did one or more reviews annually), and altogether, these reviewers carry out 1,850 reviews annually. This makes an average of 4.9 reviews per active reviewer. Average time spent per review was 7.8 hours, and the total number of hours spent on reviewing annually by the respondents amounted to 12,614 hours. The number of hours for UiT as a whole should be higher, since many researchers did not respond to the survey. While priority given to peer review is relatively high, 26% of the requests to review are declined. Few researchers have received courses and training, while some, especially new and inexperienced reviewers, would like courses and training. All over, many researchers want reviewing to become included in the performance-based budgeting system, and thought peer review should be better acknowledged in their own institution, and be more visible as a part of scientific research. This discussion shows that there are reasons **for making peer review more visible and recognized,** especially considering the extensive amount of time and work spent on reviewing.

Ethics

Peer review is related to a number of concerns relating to the ethics of the field – fears of referee theft of authors' ideas, charges of financial conflicts of interest among authors, complaints about harshly worded anonymous referee reports, concerns over scientists' lack of candour about their research misconduct, and various other items. All these fears are outlined in a review article written by Souder (Souder, 2011). The topic is very closely linked with the event organised in July 2010 in Singapore where the first global code on research integrity was agreed on. The statement was signed by science organisations around the world; it lists 14 rules for maintaining research integrity and society's trust in science. The statement covers areas such as research integrity, trustworthiness and accountability before moving on to different aspects of research, for example authorship, peer review and publication acknowledgement. The list of responsibilities includes advice on how to report irresponsible research practices and how to respond to them. The research environment and wider social considerations are also mentioned as factors that contribute to, or prevent, research misconduct (Singapore Statement, 2010)

Main findings

- **Reviewers:**
  - The 'market' for peer reviewers needs to be analysed, including the possible identification of non-financial incentives. Surveys and other evidence have shown that there are various reasons why academics participate in peer review. Not all motivations are altruistic, and there is no reason why they should be. However, a central element, without which the peer review system would not exist, is the professional commitment

to contribute to the academic public good. Each university in receipt of public funds should accept an obligation to encourage its researchers to engage in these activities, recognising that peer review is an essential part of the fabric of academic life – the costs of which are met by the funds allocated by the Funding Councils to support research infrastructure. Develop a more sophisticated understanding of the costs of peer review, and maintain efforts to ensure that the peer review processes are proportionate to the scale of award.

- Need to train the reviewers - traditionally reviewers have been 'part-and-parcel' of the system applying their expertise in the reviewing process without experiencing any training.
- The idea of drawing up a common database of "certified" experts, which was very popular at the beginning of 21th century, over time, started to be treated carefully. In fact what might appear initially simple and attractive to implement, raises a number of problems (how and by whom the certification is made; how discipline boundaries are defined; how possible reputational consequences for experts who are deemed unsuitable for the database should be dealt with).
- Different kinds of peers should be used for different purposes – specifically targeting specialists in translational or high-risk, innovative research, for example, where this is the desired outcome. This has important implications for funding bodies; since reviewers both identify and define good research, an extensive understanding of different views within a field will be required by the person selecting reviewers.

- **Peer Review judgments**
  - There is no one model that all should follow – the peer review is not in fact a single process, but rather a flexible set of mechanisms. The prerequisite of good peer review is its permanently improving character, and involvement of high level stakeholders and experts.

  - High level of expertise among the peer reviewers is certainly a must, however quality evaluations come from diverse panels of experts, which might include a mixture of backgrounds and, if relevant, different straightforward approaches and they will usually have to be tailored to the type of call. Panel composition should take into account appropriate coverage of the relevant scientific and technological domains, including interdisciplinary and socio-economic aspects. It should be also, as far as possible, balanced in terms of gender, age, affiliation and nationality, including representatives from the civil society.
  - Quality should not be sacrificed in favour of relevance and impact. Applied research ought to meet the same standards of research design, sample selection and evidential inference that applies to any sort of work (allowing for the practical difficulties of conducting applied research). Indeed, if research is being used by policy makers to take decisions on matters that have a direct effect on the quality of citizens' lives, the standards ought to be as high as possible. Similarly, novelty cannot be regarded as a substitute for quality. Ensure that considerations of applicability and relevance do not compromise judgements of quality. Set aside funds for risky, speculative projects. Ensure that there is a healthy balance between strategic funding and responsive mode projects. Encourage endowments within universities to support small grants for innovative, high risk research. Ensure that the process of selecting topics for strategic initiatives is also subject to peer review.
  - A major problem in peer-review judgments is the substantial error that occurs when a large number of reviewers evaluate only a few articles or grant applications. To

overcome this problem, it is suggested to try a 'reader system' in which a select few senior academics read all the articles or grant applications in their particular field of scholarship.

- **Improving and Modification of PRev**

- The team of LiquidPub project presents their ideas of several novel approaches for research evaluation (bidding as an indicator of importance, peerevaluation.org vs. ucount, use of community opinions, incentives to participate, the role of the internet) which are alternative or complementary to traditional peer review. The landscape of the processes for the evaluation of research outputs and of researchers are changing. In the near future we envision the growth of various tools for research evaluation, including open source and those operating with open API/protocols. Such tools would primarily operate on the Web and include the variety of methods for research evaluation, so that PC chairs or journal editors (or even people playing some new emerging roles which do not exist yet) will be able to choose. Examples of tools with such functionalities already emerge (e.g., Mendeley, Peerevaluation.org, Interdisciplines) but it is not yet clear how these tools can be connected and which of them will be adopted widely enough to have a normative effect. Attention should be paid less to designing "the" scientific evaluation system of tomorrow – something that, like "the" peer review process, will be an emergent phenomenon based on the different needs of different disciplines and communities. Instead, attention should focus on ensuring interoperability and diversity among the many possible tools that scientific evaluation can make use of.
- Potential modifications to the grant peer review process may be considered to improve efficiency or effectiveness. With respect to efficiency, for example, improvements could be brought about by *moderating demand* to ensure that the number of applications received is kept below a certain threshold – thus reducing the burden on reviewers and applicants. This could be achieved by (i) reducing advertising; (ii) changing deadline systems for funders that use fixed milestones for submission; or (iii) limiting the number of applications from particular institutions. It may also be possible to *streamline assessment procedures* using tighter systems of triage on applications received.
- Other potential cost-saving measures include (1) reducing the number of external referees involved in peer review of grant applications, and (2) increasing the use of technology – including videoconferencing – so that peer review panellists do not have to gather in one place for scoring meetings.
- Improving the capacity of peer review to support applied research**:** the selections of the Panel members from both academic peer review and decision making constituencies; but educators and communication experts may also participate if the proposal in question is likely to be a high-impact area of research. The aim is thus to evaluate research proposals both in terms of their scientific merit and the potential impact they may have.
- Improving the capacity of peer review to support innovative research, s.c DARPA model: a narrowed down version of peer review, in which there is no panel, simply 'expert' judgement by a specially selected programme manager.

2. Survey on Peer Review Practices

The biggest concern expressed about PRev is the objectivity of assessment concerning gender, nationality, major field of study and institutional affiliation. Respondents' affiliation and nationality geography shows a rather interesting picture – while respondents belong to 79 countries by nationality, their places of affiliation are in 66 countries. The lion's share of respondents belong to fifteen countries (their share by nationality is 76%, and the share by affiliation is 81,2%). As we see in Figure 3, the centres of complete attraction are the United Kingdom and Switzerland.

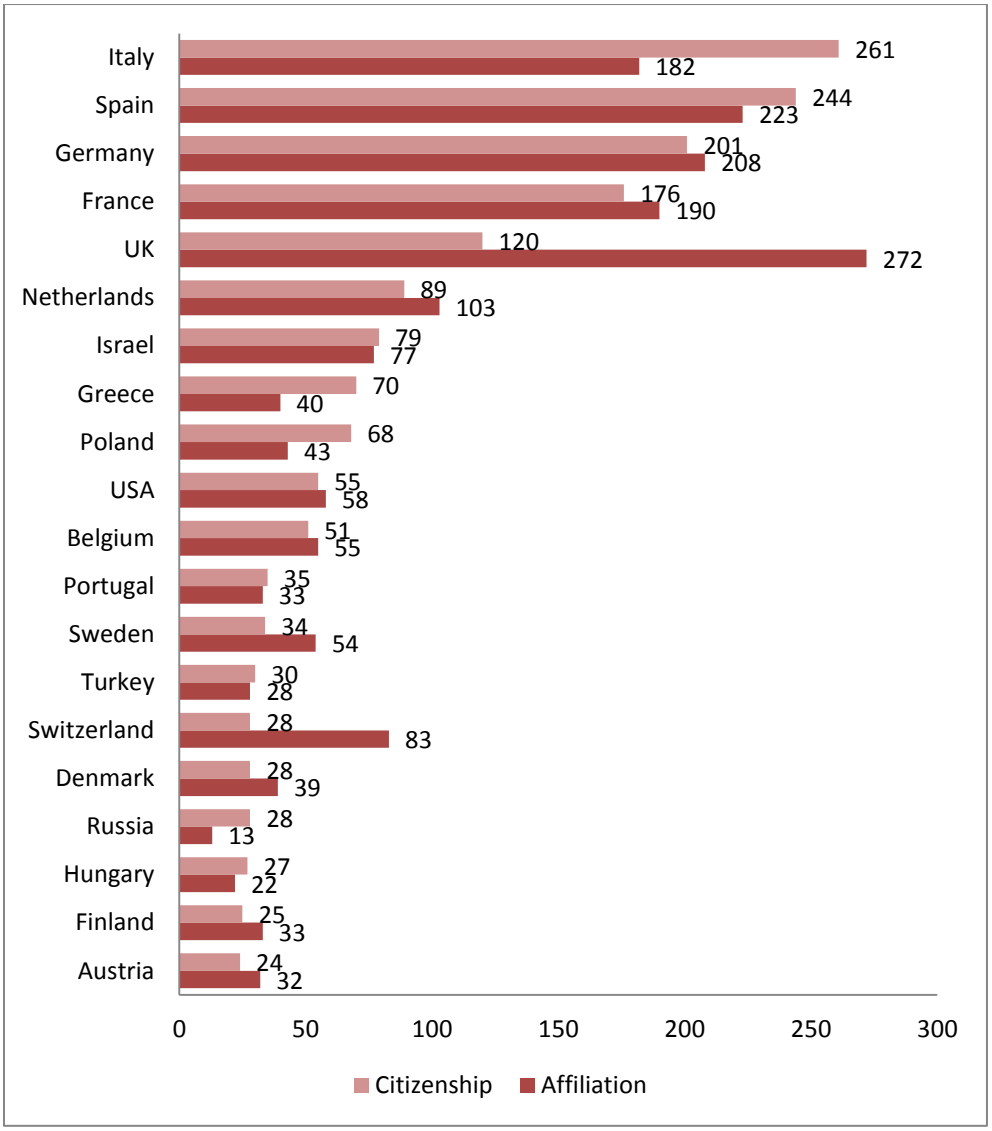

Figure 3. The proportions of respondents by citizenship and affiliation (by number)

As we see in Figure 4, the largest numbers of respondents belong to natural sciences, however the overall proportions of fields match the corresponding figures of the European Union (Eurostat, 2009 Tables 3.11 and 3.12).

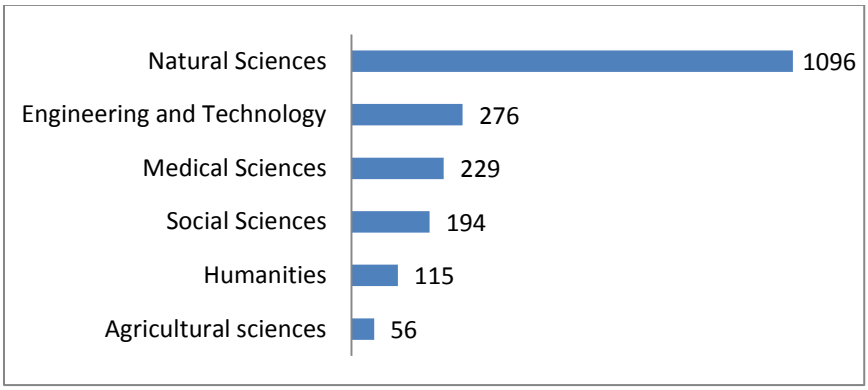

Figure 4. The number of respondents by field

Among respondents, 63.25% (n = 1337) were men, and 30.8% (n = 652) were women (125 did not specify their gender). This is in conformity with the general practice in Europe, on an average, 30% of European researchers are women (She figures, 2009). As we see in Figure 5, there are some differences between fields. Traditionally women have a higher proportion in Humanities and Social Sciences.

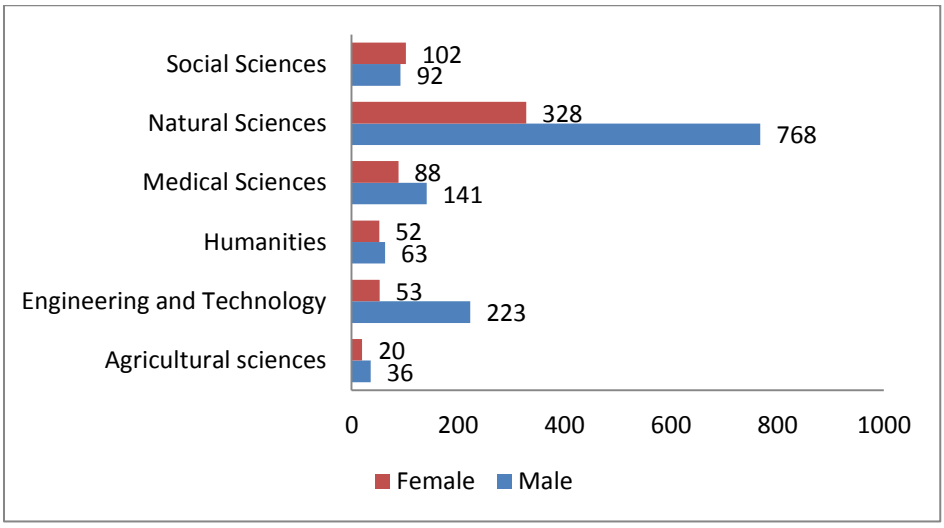

Figure 5. The proportion of respondents by field of science and technology, by gender

Figure 6 again reflects the situation prevalent in Europe – the proportion of women decreases with their career progression. Only 15% of full professors in European universities are women, and women are under-represented on scientific decision-making boards in almost all European countries (She figures, 2009). The proportion of women professors in this survey is significantly higher.

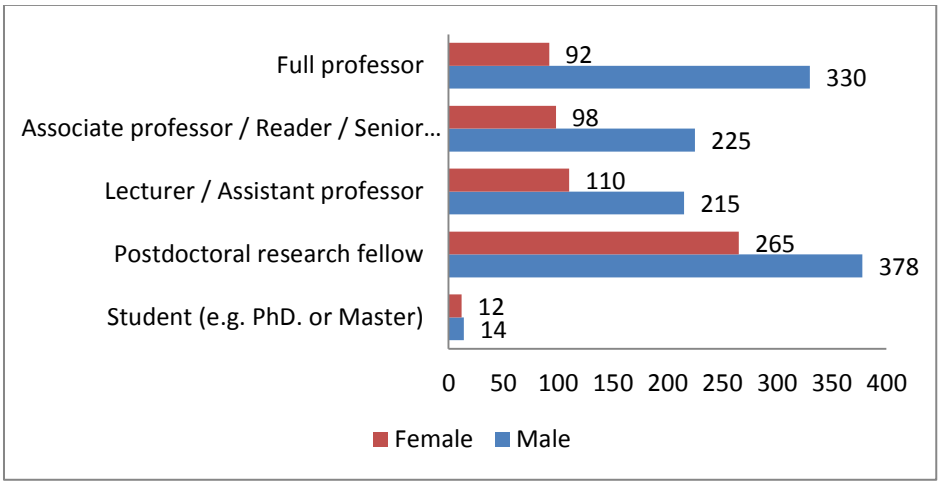

Figure 6. The proportion of respondents by gender and by academic position

Synopsis:

- Respondents' affiliation and nationality geography illustrates the ongoing trends in the mobility of researchers in Europe and in the world. While respondents belong by nationality to 79 countries, their places of affiliation are in 66 countries.
- The survey reflects the situation prevalent in Europe – the proportion of women decreases with their career progression.
- The largest numbers of respondents belong to natural sciences and that may influence the results, however the overall proportions of fields match the corresponding figures of the European Union.

Experience in the PRev process (as applicant or reviewer)

Among all respondents, 1975 defined themselves as applicants and 889 as also reviewers. 67.3% of applicants and 71.4% of reviewers were men. By academic position, postdoctoral research fellows (37%) and full professors (25%) constituted the largest share of applicants. Among reviewers, full professors (41%) and associated professors (24%) dominated.

1.2. Respondents' experience as an applicant

We wanted to know in the survey whether there are any differences between applications on national, international or supranational levels. The results obtained showed that the lion's share of respondents on all levels have made less than ten applications during their scientific career. Differences between the levels can be detected in case of applicants who have made dozens of applications – here dominance of projects of the national level is apparent. Hundreds of applicants had been made on the national level by 14 applicants, on the international level by three applicants, and on the supranational level by one applicant. Among the three applicants who had been so active on the international level, two had never applied on the national level.

Table 1. The proportion of applications on different levels

%	Personal Grant or Fellowship				As coordinator of a project			
	None	Fewer than 10	11-100	More than 100	None	Fewer than 10	11-100	More than 100
National	13,4	62,2	23,7	0,8	37	48,4	14,5	0,2
International	30,6	59,2	10	0,2	58,2	38,5	3,3	0
Supranational	16,8	77,2	5,9	0,1	54,2	44,2	1,6	0

According to the estimates of respondents, the success rate of their applications has been extremely high: the success rate of 67.4% of national applications, 60.2% of international applications and 70.1% of supranational applications was higher than 50%. The average success rate of personal grant or fellowship applications on the national level is 56.75%, on the international level – 52.85%, and on the supranational level – 62.9%.

The level is lower in case of coordination: in case of national applications the average success rate is 46.15%, in case of international projects – 40.95%, and in case of supranational projects – 43.8%.

Slight differences can be observed between the various academic positions. Full professors are apparently with a higher success rate in obtaining projects on the national level (Figures 9 and 10). They are, however, with the lowest success rate in obtaining personal grants on the supranational level.

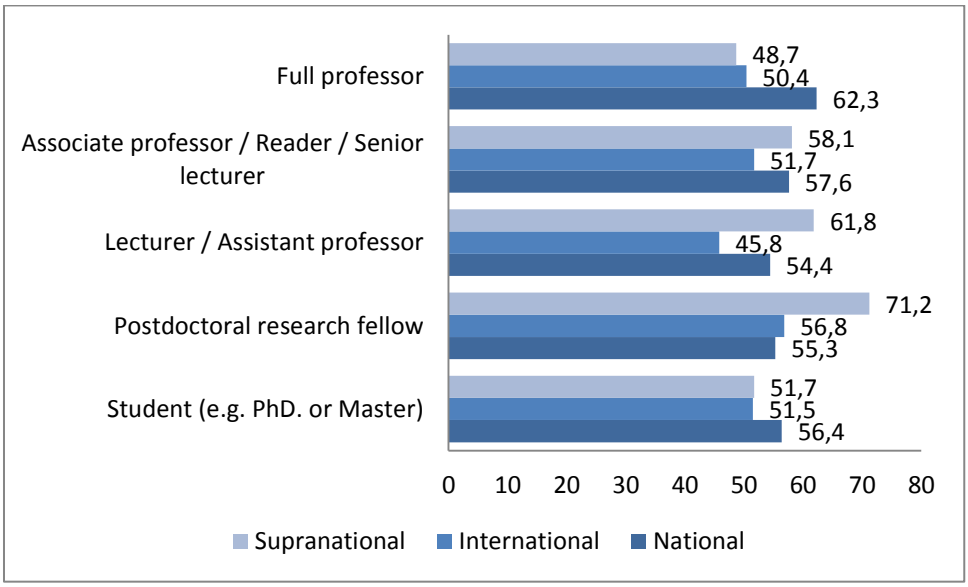

Figure 7. The average success rate of personal grant or fellowship applications by field

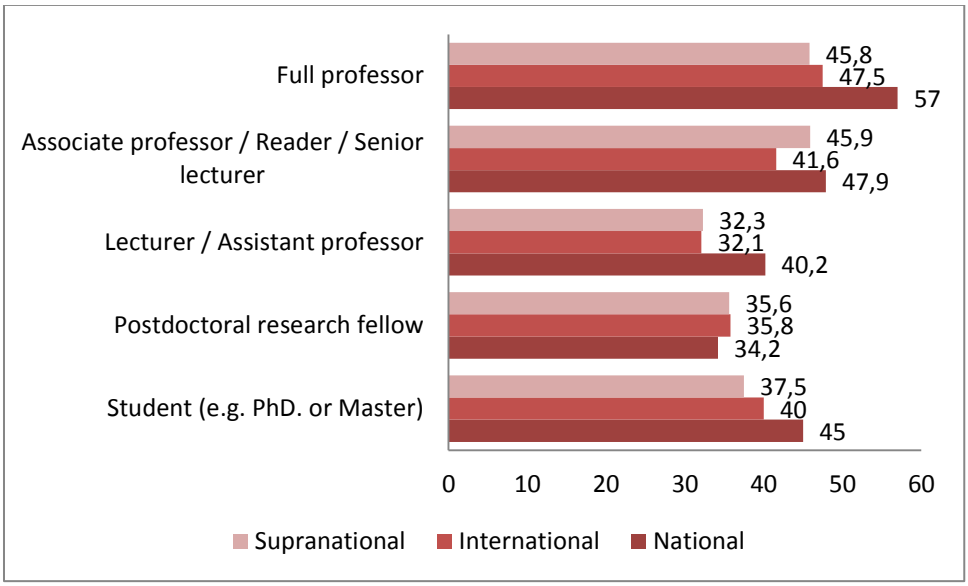

Figure 8. The average success rate of coordination by field

The main reasons for rejection of applications are the same on all three levels – lack of funding (61.4%), and remaining under the evaluation threshold (49%).

In the majority of cases applicants were informed of the reasons for rejection, only 8% of respondents had not received any responses.

Synopsis

- The success rate of the respondents' applications has been extremely high: 67.4% of national applications, 60.2% of international applications and 70.1% of supranational applications have a success rate higher than 50%.
- Some differences can be observed between the various academic positions, full professors are apparently with a higher success rate in obtaining projects on the national level, they are, however, with the lowest success rate in obtaining personal grants on the supranational level.
- The main reasons for rejection are lack of funding (61.4%), and remaining under the evaluation threshold (49%).

Respondents' experience as a reviewer

From the total number of respondents, 43.4% stated that they had experience as reviewer. As we already stated, the lion's share of respondents belong to Natural Sciences. In case of reviewers we can notice a slight increase in other fields at the expense of the Natural Sciences (Figure 9).

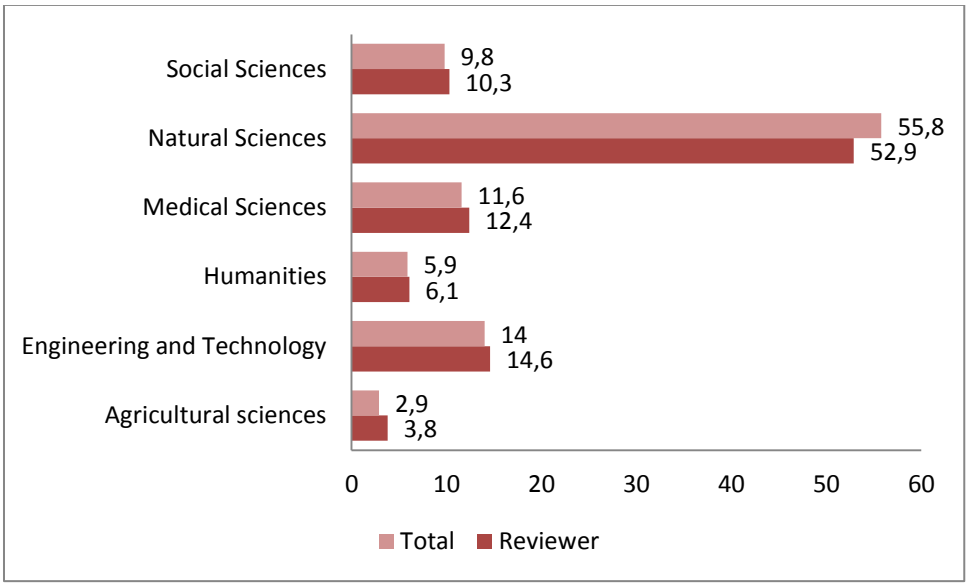

Figure 9. The proportion of reviewers by field (%)

We wanted to know in the survey why the respondents were interested in participation in the Peer Review process. In the majority of cases (Figure 10) it was because of wider ethical considerations, such as "*I feel a general obligation towards the field*", „*I want to ensure the quality of my field*", „*I want to educate/help fellow researchers*", or because of self-education purposes – "*It gives me an overview of my field*". However, the fact that for nearly a quarter of the respondents, being reviewer was also considered as a tool to improve their reputation should not be overlooked.

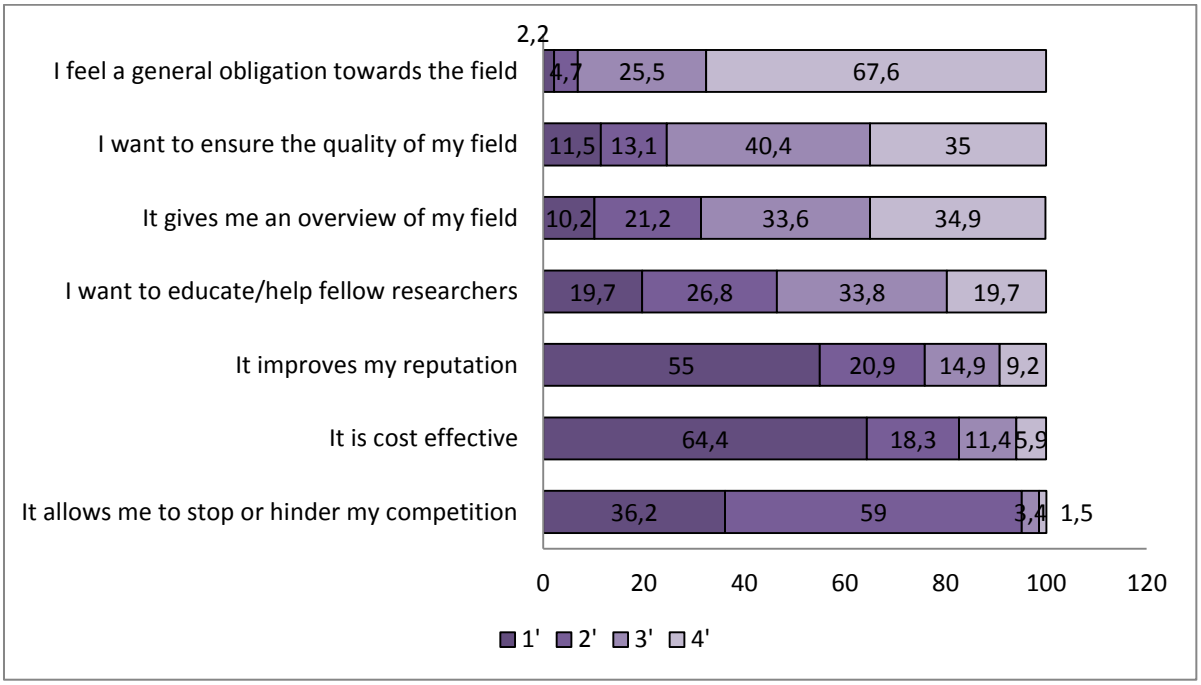

Figure 10. What were the main reasons to be a reviewer (last five years) 1 = very rarely, 2 = rarely, 3 = often, 4 = very often (in %)

Answers to the question "*Have you refused a request to review an application (in the past five years)?"* varied very much. This is especially evident in comparisons between fields and gender. While 67.3% of postdoctoral research fellows and 60.6% of lecturers and assistant professors stated that they had never refused reviewing, 57.6% of associated professors and 83.6% of full professors admitted that they had refused reviewing.

There is a slight difference between genders – 46.2% of female respondents and 37.5% of male respondents had never refused reviewing.

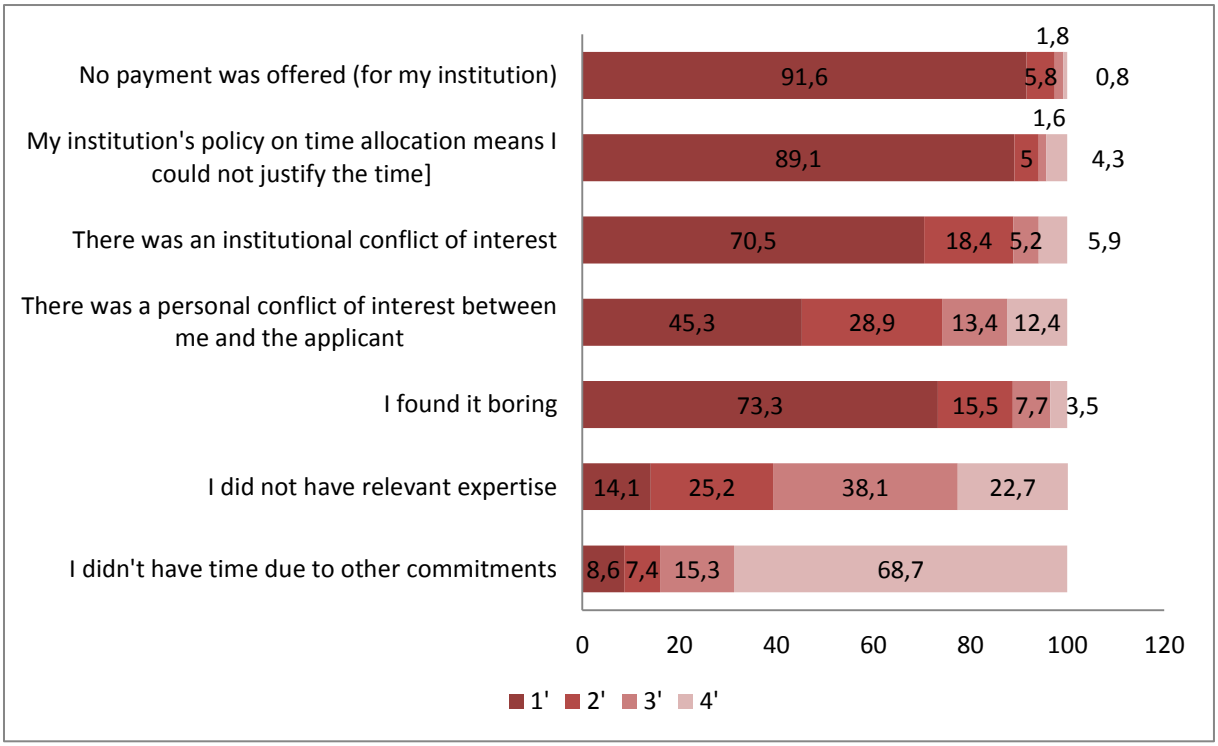

Figure 11. The main reasons for refusal to review (% of often (3) and very often (4))

The most common reasons for refusing to review were the lack of time – 84% of all respondents stated that this is often or even very often a problem. The same trends are evident in other surveys. Tite and Schroter found in their survey that the most highly rated factor important in the decision to decline to review was a conflict with other workload. Most respondents agreed that financial incentives would not be effective when time constraints are prohibitive. Reviewers were also more likely to accept to review when it was relevant to their area of interest. Reviewing should be formally recognised by academic institutions (Tite&Schroter 2007). The same outcomes appeared also in this survey (see §5).

Another reason to refuse reviewing was the feeling that they lacked the relevant expertise (60.8%). Here postdoctoral research fellows were the most confident, only 45.2% of them considered it a frequent reason, for 60.9% of full professors it was a problem (Figure 12).

It seems that in Peer Review practice it is not common to inform reviewers about the final results of applications reviewed by them. About 34% of reviewers have never been informed about results, and 23% of them have been informed very rarely.

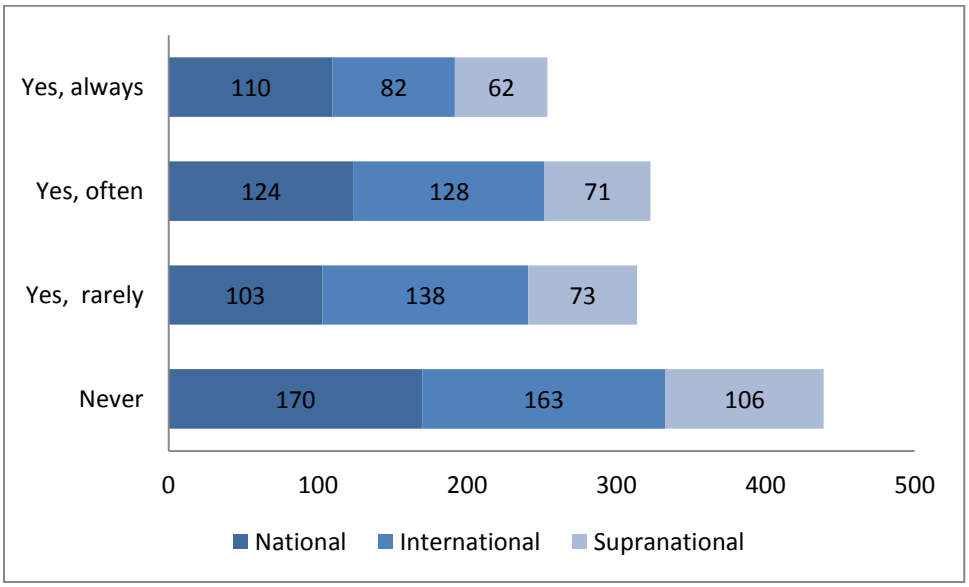

Figure 12. **Have you been informed of the final results of the application reviewed by you?**

The majority (75.3%) of respondents did not consider it necessary to receive feedback about the final results.

Synopsis

- Reasons why the respondents agreed to be reviewers were mainly linked to research ethics – obligation towards the field, intention to ensure the quality of the field, and desire to help fellow researchers. Also self-improvement was important – to receive an overview of own field.
- While the majority of postdoctoral research fellows (67.3%) and lecturers and assistant professors (60.6%) had never refused to be a reviewer, the majority of associated professors (57.6%) and full professors (83.6%) had refused reviewing.
- The most common reasons to refuse to review were the lack of time – 84% of all respondents stated that this was often or even very often a problem, the second by importance was the feeling that they lacked the relevant expertise (60.8%). Here postdoctoral research fellows were the most confident, only 45.2% of them considered it a frequent reason (for 60.9% of full professors it was a problem).
- It seems that in Peer Review practice it is not common to inform reviewers about the final results of applications reviewed by them. About 34% of reviewers had never been informed about results, and 23% of them had been informed very rarely. The majority (75.3%) of respondents did not consider it necessary to receive feedback about the final results.

S&T indicators

The quality of research output is measured with different S&T indicators. They measure the various components of research activity, including inputs, process, outputs, outcomes and impact and benefits (Assessing, 2010). Some research assessments assign different weightings or values to the various indicators. In this way, some components of research activity are valued more highly than other activities**.** There is the eternal question of whether the indicators really do measure what they are intended to measure because people adjust to the indicator value system by optimizing their indicator rather than their performance. Thus, no matter how sophisticated an indicator may be, it will never be a proof against manipulation. (Retzer&Jurasinski, 2009)

In the survey we asked respondents to assess on a five-point scale the different indicator used in research assessments. The highest rating was given to the following indicators: high ranked publications (4.5), citations (3.9), research collaborations and partnership (3.8), reputation and esteem – position as journal editors, membership of editorial boards and scientific committees and membership in learned academies (3.6), and number of prestigious national and international awards and prizes (3.6). In ratings there were no differences between men and women. But there were significant differences by fields (see Table 1). This applies in particular to humanities. Although publications received the highest scores, it is clear that understanding what a high ranked publication is varies between fields.

> "*The social sciences and humanities should have a separate evaluation system. Basing grants on ISI-ranked publications, for instance, is useless (and gives a totally wrong image of the field and the importance of the assessed work*)", and a humanities researcher from Australia adds: "*In the arts disciplines, publication record remains the key indicator of excellence though increasingly publication needs to be understood in terms of public outputs, including exhibitions, performances and so forth*". (Social sciences, Spain)

Such indicators as citations, h-index and patents are not relevant to the humanities. Estimates of indicators related to the PhD (number of PhD graduates and completion rates for graduates, number of PhD students, and employability of PhD graduates in private sector) are remarkably low.

> "*Number of PhD students is a traditional measure very common in the EU. It is utterly irrelevant, even distorting, because it creates an incentive to have as many PhD students as possible, no matter how poorly supervised. What matters is the success rate in placing PhDs in full-term, permanent employment upon completion. This reflects the actual training and mentoring provided."*(Social sciences, Germany)

Table 2. The rating of the S&T indicators by fields

Indicators	Average	Agricultural sciences	Humanities	Medical Sciences	Social Sciences	Engineering and Technology	Natural Sciences
Effect indicators							
Citations	3.9	4.3	2.9	4.1	3.8	3.8	4.0
H index	3.5	3.8	2.4	3.6	3.4	3.5	3.6
Number of prestigious national and international awards and prizes	3.6	3.7	3.4	3.4	3.6	3.7	3.7
Employability of PhD graduates (in private sector)	2.8	3.0	2.1	2.8	2.4	3.3	2.8
Input indicators							
Recruitment of PhD students	3	3.2	2.9	3.0	2.6	3.2	3.0
External funding	3.5	3.7	3.2	3.6	3.4	3.5	3.5
Structure indicators							
Number of PhD students	2.9	3.2	2.6	3.0	2.6	3.2	2.9
Research collaborations and partnership	3.8	4.1	3.7	3.8	3.7	4.0	3.7
Reputation and esteem	3.6	3.9	3.6	3.5	3.8	3.7	3.5
Output indicators							

Publications	4.5	4.5	4.6	4.6	4.5	4.3	4.5
Non-bibliographical outputs	2.8	3.1	3.0	2.8	2.8	2.9	2.8
Number of PhD graduates and completion rates for graduates	3.2	3.5	2.7	3.3	3.0	3.3	3.3
Patent development	2.9	3.1	2.4	2.9	2.7	3.2	2.8
Public outreach	3.2	3.8	3.0	3.4	3.5	3.4	3.1
Social indicators							
Relevance to citizens' concerns	3	3.3	2.7	3.5	3.4	3.1	2.8
Relevance to global societal challenges	3.2	3.7	2.9	3.6	3.6	3.3	3.0
Usefulness to policy decision makers	2.9	3.7	2.5	3.3	3.3	2.9	2.7
Contributing to science education	3.5	3.9	3.3	3.7	3.5	3.5	3.5
Relevance to science communication initiatives	3.2	3.8	2.8	3.5	3.3	3.2	3.1
Process indicators							
Seminar and conference activity	3.3	3.2	3.6	3.1	3.3	3.4	3.2
Invited keynotes	3.5	3.7	3.3	3.4	3.3	3.6	3.5
International visiting research appointments	3.3	3.4	3.6	3.2	3.3	3.4	3.2

Views vary also at different career stages. Particularly different are the preferences of students and professors. While relevance to global societal challenges, public outreach, contributing to science education, usefulness to policy decision makers, relevance to citizens' concerns received the highest ranks from students, professors ranked the same indicators the lowest. And vice versa, while high ranked publications, citations, number of prestigious national and international awards and prizes, h-index received the highest ranks from professors, students ranked the same indicators the lowest.

The overall position was that various indicators should be weighted differently (agreed by 66.4%) and that there is a need to have different weightings for the various indicators in different subject areas (agreed by 68.5%) as well as a need to have various indicators or different weighting for the various indicators at different career stages (agreed by 69.1%).

This leads to the conclusion that there is no such thing as an objective indicator because indicators are rarely a direct measurement. The choice of indicators and weightings is vital. They must be fit-for-purpose, appropriate and verifiable. (Assessment, 2010, 36)

Synopsis:

- The most favoured indicator was high ranked publications (4.5),and this is the only one which was accepted by researchers from all fields.
- Other favoured indicators were citations (3.9), research collaborations and partnership (3.8), reputation and esteem – position as journal editor, membership of editorial boards and scientific committees and membership in learned academies (3.6), and number of prestigious national and international awards and prizes (3.6).
- The overall position was that various indicators should be weighted differently (agreed by 66.4%) and that there is a need to have different weightings for the various indicators in different subject areas (agreed by 68.5%) as well as a need to have various indicators or different weighting for the various indicators at different career stages (agreed by 69.1%).
- Indicators have to be fit-for-purpose, appropriate and verifiable.

Criticism of Peer Review

Conflict of interest, gender issues, subject biases are the main concerns about PRev shortcomings. We defined eleven different biases in this survey and asked respondents to rank them on a five point scale. It was surprising that there was an almost complete consensus among the respondents, both in terms of gender, field as well as the academic position. The most urgent concern was related to the so-called Matthew effect – "to those who have, more shall be given"(rating 3.8), institutional bonus (3.6), friendship bonus (3.6). It seems that these were the most pressing ones in medical sciences and social sciences. At the same time, it was surprising that gender issues were completely in the background (Table 2). Although language does not seem to be a particular problem, the literature review on the topic of PRev shows a slightly different picture (https://www.archimedes.ee/acumen/). Analysing the proportion of papers by the authors' geography of affiliation which relates to the languages[i], the most commonly encountered topics in the papers were: PRev best practice; the pros and cons of PRev versus bibliometrics; bias (of all forms); evaluation of quality research / research assessment evaluation (which is UK specific); effectiveness of research grant/applications; PRev in teaching; teaching PRev to potential reviewers; ethical issues; difficulties facing researchers in non-English speaking countries; overviews of PRev; PRev challenges of interdisciplinary research & innovative research. When we analysed the ten issues of concern most written about by English-speaking authors and English as a second language speaking authors, we found an unexpected result of this analysis concerning the ranking of 'bias' related papers. In case of English speaking nations the issue of bias was ranked eighth, in the case of non-English speaking countries it was ranked second (in case of EU-27 non-English speaking countries even first).

> *„Biases can become more prevalent if reviewers have little time to assess application. It is likely that in these cases, reviewers will rely more on fast and frugal heuristics, and on stereotypes (i.e. relying more on the institutional affiliations, the gender, the reputation of the applicant, and so on). If we are speaking about fairness here, the "Matthew effect" is a major problem."*(Social sciences, France)

There is an everlasting question: does the PRev review fulfill its declared objective of identifying the best science and the best scientists? Bornmann et al (Bornmann, 2008) analysed the Long-Term Fellowship and the Young Investigator programmes of the European Molecular Biology Organization. The statistical analyses pointed out that between 26% and 48% of the decisions made to award or reject an application showed an error. Even though for a part of the applicants, the selection committee did not correctly estimate the applicant's future performance, the results show a statistically significant correlation between selection decisions and the applicants' scientific achievements, if quantity and impact of research publications are used as a criterion for scientific achievement.

Synopsis

Among eleven different biases which may occur in the PRev process, the most urgent concern was related to the so-called Matthew effect – "to those who have, more shall be given"(rating 3.8), institutional bonus (3.6), friendship bonus (3.6), and entrenched academic traditionalism (3.5). There was an almost complete consensus among the respondents, both in terms of gender, field, as well as the academic position.

Table 3. The rating of the selection of bias by field

Bias	Average	Agricultural sciences	Engineering and Technology	Humanities	Medical Sciences	Natural Sciences	Social Sciences
Institutional bonus	3.6	3.6	3.5	3.5	3.7	3.6	3,8
Geographical origin	3.1	3.1	2.9	3.0	3.2	3.1	3.2
Language (favouring of papers written in English)	3.1	3.2	3.3	3.4	3.3	3.0	3.5
Conflict of interest	3.1	3.0	3.2	3.1	3.3	3.1	3.2
Gender	2	1.8	1.8	2.4	2.0	2.0	2.3
The scope of the research	3.1	2.9	3.1	3.0	3.1	3.1	3.0
Peer Review as a time consuming process	3.1	3.0	3.1	3.3	3.2	3.1	3.2
High cost of Peer Review	2.6	2.7	2.7	2.5	2.7	2.6	2.9
Matthew effect	3.8	3.7	3.6	3.5	4.0	3.8	4.0
So-called friendship bonus	3.6	3.3	3.5	3.4	3.8	3.6	3.8
Entrenched academic traditionalism	3.5	3.5	3.4	3.6	3.8	3.5	3.8

The future of the peer review system

How to make PRev more fair and objective? The majority of recommendations were related to the person – the reviewer. The overwhelming view was that the people who agree to participate in the PR process should be recognized.

> *I think reviewer work should be more appreciated and count more at the institutions where we work. Thus, information on who reviews for granting body should be made public just like information about editorial boards of journals. When the institutions understand that reviewing grant proposals is the same creation of an academic culture as reviewing for journals then reviewers will be doing a better job reviewing because they will want to be on such bodies.* (Social sciences, Poland)

> *Peer review will soon become a full-time task at least for a few month period during the year. Rewarding in term of visibility will become mandatory as the positive evaluation of a good record of reviewing activity in case of job application.(Physics, UK)*

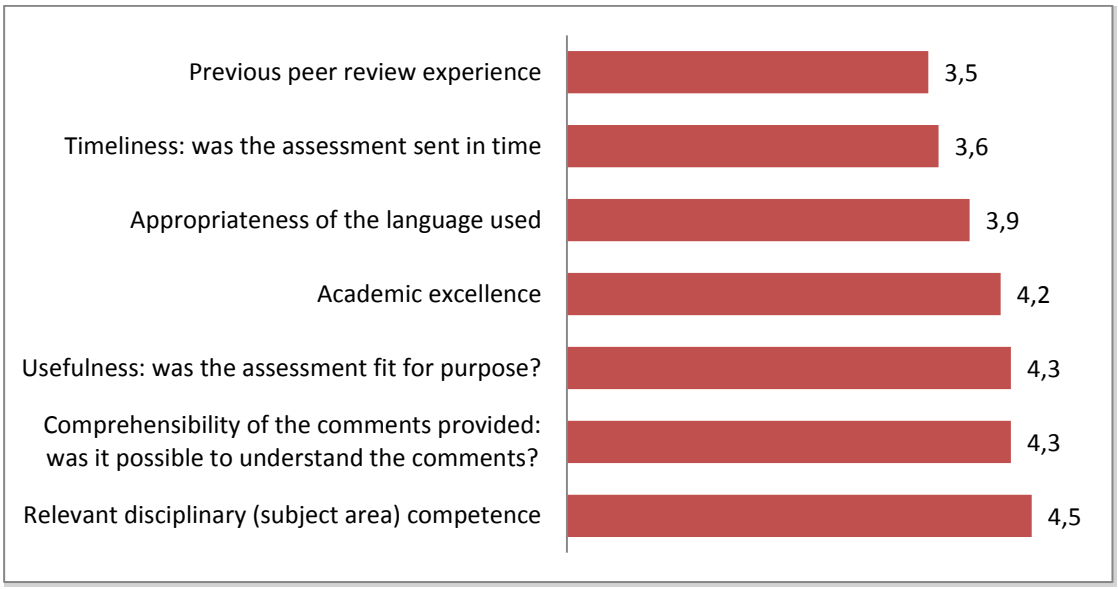

Figure 14. The rating of criteria essential for a good reviewer.

As we see in Figure 14, all criteria essential for a good reviewer are highly valued. A good reviewer has relevant disciplinary competence and academic excellence, the comments are comprehensive and useful, the review is written in appropriate language and it is submitted in time, previous peer review experience is also needed.
Sometimes the PRev process has been accused of anonymity.

> *"It should not be anonymous any more. Too many referees hide behind anonymity, and write ever nastier things." (*Mathematics, Austria*)*

In the survey we wanted to know to which extent respondents want to receive information about reviewers. As we see from Figure 16, the primary interest is in having a reviewer's written evaluation available to the applicant, excluding reviewer's name (54.6%), and that an applicant should have the possibility to read and respond to the reviewer's comment before the final decision (49.9%). A survey conducted in organizations within the ESF show that 47% of the participating organisations allow applicants to reply to the assessment of their proposals in all or some instruments, while 53% generally do not allow any replies (ESF Survey Annex, 2011). It seems that respondents valued also the idea that a reviewer's rating (grade) should be available to the applicant (37.6%). It seems to be an issue especially in medical sciences – 40.6% of respondents favoured this idea. RAND Europe's Health Research System Observatory report (Ismail, et al, 2009) showed that the need for chanes and improvements is especially spread in medical sciences community.

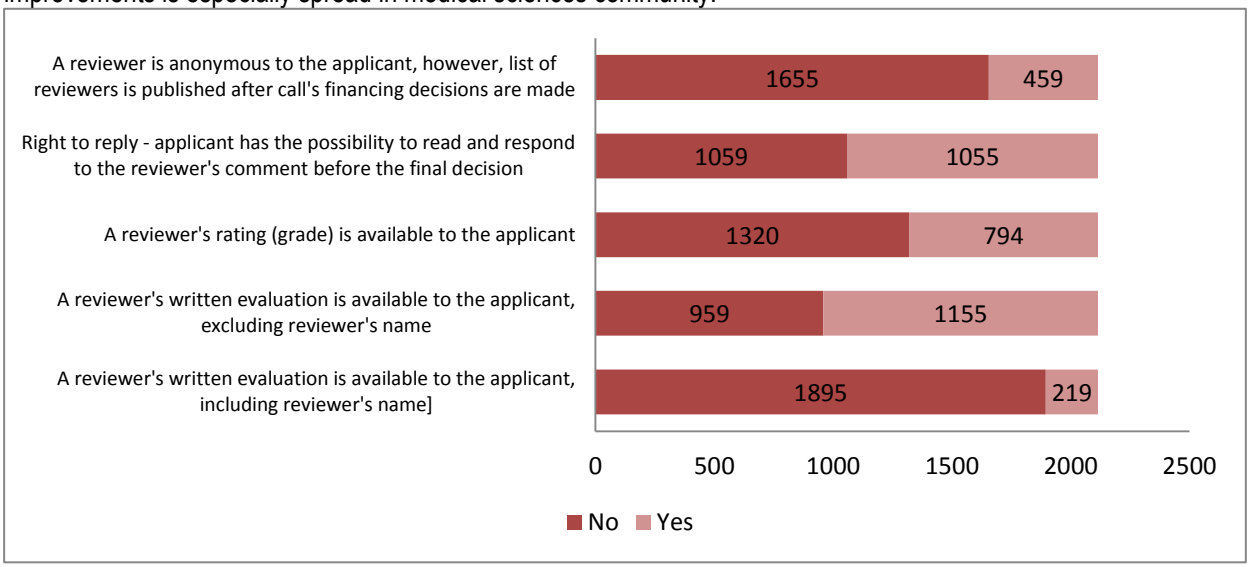

Figure 15. The extent of information about reviewers which should be available to the applicants (number)

More than a half (56%) of respondents had had experience with a system that allows applicants to nominate possible reviewers, and the majority of them (88.5%) had used this possibility.

A little less (48.3%) known is a system that allows applicants to exclude reviewers but the same time majority (56.9%) of those who know the system have used it. Full professors have been particularly active here – 63.2% have used this possibility.

The majority of respondents (59.9%) favour the opinion that there is a need for improvements. This seems to be particularly relevant in medical sciences where 72.4% of respondents voted for changes.

Synopsis

- The majority of the proposals to improve the Peer Review system were related to reviewers. The overwhelming view was that the people who agree to participate in the PR process should be recognized. A good reviewer has relevant disciplinary competence and academic excellence, the comments are comprehensive and useful, the review is written in appropriate language and it is submitted in time, previous peer review experience is also needed.
- The majority of respondents wanted to have a reviewer's written evaluation available to the applicant, excluding reviewer's name (54.6%), and considered that an applicant should have the possibility to read and respond to the reviewer's comment before the final decision (49.9%).

Main findings:

- Reasons why the respondents agreed to be reviewers were mainly linked to research ethics – obligation towards the field, intention to ensure the quality of the field, and desire to help fellow researchers. Also self-improvement was important – to receive an overview of own field.
- While the majority of postdoctoral research fellows (67.3%) and lecturers and assistant professors (60.6%) had never refused to be a reviewer, the majority of associated professors (57.6%) and full professors (83.6%) had refused reviewing.
- The most common reasons to refuse to review were the lack of time – 84% of all respondents stated that this was often or even very often a problem, the second by importance was the feeling that they lacked the relevant expertise (60.8%). Here postdoctoral research fellows were the most confident, only 45.2% of them considered it a frequent reason (for 60.9% of full professors it was a problem).
- It seems that in Peer Review practice it is not common to inform reviewers about the final results of applications reviewed by them. About 34% of reviewers had never been informed about results, and 23% of them had been informed very rarely. The majority (75.3%) of respondents did not consider it necessary to receive feedback about the final results.
- More than a half (56%) of respondents had had experience with a system that allows applicants to nominate possible reviewers, and the majority of them (88.5%) had used this possibility.
- A little less (48.3%) known is a system that allows applicants to exclude reviewers but the same time majority (56.9%) of those who know the system have used it. Full professors have been particularly active here – 63.2% have used this possibility.
- The majority of respondents (59.9%) favour the opinion that there is a need for improvements. This seems to be particularly relevant in medical sciences where 72.4% of respondents voted for changes.
- The majority of the proposals to improve the Peer Review system were related to reviewers. The overwhelming view was that the people who agree to participate in the PR process should be recognized. A good reviewer has relevant disciplinary competence and academic excellence, the comments are comprehensive and useful, the review is written in appropriate language and it is submitted in time, previous peer review experience is also needed.
- The majority of respondents wanted to have a reviewer's written evaluation available to the applicant, excluding reviewer's name (54.6%), and considered that an applicant should have the possibility to read and respond to the reviewer's comment before the final decision (49.9%).

References

Abramo, Giovanni; D'Angelo, Ciriaco Andrea (2009) A Decision Support System for Public Research Organizations Participating in National Research Assessment Exercises. JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 60 (10), 2095-2106

Adams, J (2009) The use of bibliometrics to measure research quality in UK higher education institutions. ARCHIVUM IMMUNOLOGIAE ET THERAPIAE EXPERIMENTALIS, 57 (1), 19-32

Aksnes, D, Rorstad, K, Piro, F, Sivertsen, G (2011) Are Female Researchers Less Cited? A Large-Scale Study of Norwegian Scientists. JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 62 (4), 628-636

Allen, Liz; Jones, Ceri; Dolby, Kevin; Lynn, David; Walport, Mark (2009) Looking for Landmarks: The Role of Expert Review and Bibliometric Analysis in Evaluating Scientific Publication Outputs. PLOS ONE, 4 (6)

Atkinson, M. (2001) 'Peer Review' culture. SCIENCE AND ENGINEERING ETHICS, 7(2), 193-204

Auspurg, K; Hinz, T; Guedler, J (2008) Emergence of an Academic Elite? KOLNER ZEITSCHRIFT FUR SOZIOLOGIE UND SOZIALPSYCHOLOGIE, 60 (4), 653-

Auspurg, K; Hinz, T; Guedler, J (2009) Ghost's debate or what is sociological clarification? KOLNER ZEITSCHRIFT FUR SOZIOLOGIE UND SOZIALPSYCHOLOGIE, 61 (3), 469-474

Barre, R; Regibeau, P; Lepori, B; Aguillo, I; Siedschlag, J; Soboll, H; Tubbs, M; Veugelers, R; Ziarko, W. (2009) ERA Indicators and Monitoring Expert Group Report October 2009. Luxembourg: Publications Office of the European Union

Bazeley, P (1998) Peer review and panel decisions in the assessment of Australian Research Council project grant applicants: what counts in a highly competitive context? HIGHER EDUCATION, 35 (4), 435-451

Birukou, A; Wakeling, J; Bartolini, C; Casati, F; Marchese, M; Mirylenka, K; Osman, N; Ragone, A; Sierra, C; Wassef, A (2011) Alternatives to peer review: novel approaches for research evaluation. COMPUTATIONAL NEUROSCIENCE, 5 (56), DOI: 10.3389/fncom.2011.00056

Blackburn, JL; Hakel, MD (2006) An examination of sources of peer-review bias. PSYCHOLOGICAL SCIENCE, 17 (5), pp.378-382

Bloch, S; Walter, G (2001) The Impact Factor: Time for change. AUSTRALIAN AND NEW ZEALAND JOURNAL OF PSYCHIATRY, 35 (5), 563-568

Bobe B, Viala H (1997) Panel of experts of the European Commission. In, Callon M, Laréro P, Mustar P. (eds), The strategic management of research and technology: Evaluation of programmes. Paris: Economica International.

BONUS (2006) Guidelines for a Common Evaluation Scheme for a Joint Baltic Sea

Bordage, G. (2001) Reasons Reviewers Reject and Accept Manuscripts: The Strengths and Weaknesses in Medical Education Reports. ACADEMIC MEDICINE, 76(9), 889-896

Bornmann L., Daniel H. (2005). Criteria Used by a Peer Review Committee for Selection of Research Fellows A Boolean Probit Analysis. INTERNATIONAL JOURNAL OF SELECTION AND ASSESSMENT, 13 (4), 296-303.

Bornmann, L.; Wallon, G.; Ledin, A.(2008). Does the Committee Peer Review Select the Best Applicants for Funding? An Investigation of the Selection Process for Two European Molecular Biology Organization Programmes. PLOS ONE, 3 (10)

Bornmann, L; Daniel, HD (2005) Selection of research fellowship recipients by committee peer review. Reliability, fairness and predictive validity of Board of Trustees' decisions. SCIENTOMETRICS, 63 (2), 297-320

Bornmann, L; Daniel, HD (2007) Convergent validation of peer review decisions using the h index - Extent of and reasons for type I and type II errors. JOURNAL OF INFORMETRICS, 1 (3), 204-213

Bornmann, L; Mutz, R; Daniel, HD (2007) Gender differences in grant peer review: A meta-analysis. JOURNAL OF INFORMETRICS, 1 (3), 226-238

Bornmann, L; Mutz, R; Daniel, HD (2008) How to detect indications of potential sources of bias in peer review: A generalized latent variable modelling approach exemplified by a gender study. JOURNAL OF INFORMETRICS, 2 (4), 280-287

Bornmann, L; Mutz, R; Daniel, HD (2009) The influence of the applicants' gender on the modeling of a peer review process by using latent Markov models. SCIENTOMETRICS, 81 (2), 407-411

Brinn, T; Jones, MJ; Pendlebury, M (2000) Measuring research quality: peer review 1, citation indices 0. OMEGA-INTERNATIONAL JOURNAL OF MANAGEMENT SCIENCE, 28 (2), 237-239.

British Academy (2007) Peer Review: The Challenges for the Humanities and Social Sciences, available at: http://www.britac.ac.uk/policy/peer-review.cfm.
Brown, T (2004) Peer Review and the acceptance of new scientific ideas. Discussion paper from a Working Party on equipping the public with an understanding of peer review November 2002 – May 2004. Published by Sense About Science. http://www.senseaboutscience.org
Canibano, C; Otamendi, J; Andujar, I (2009) An assessment of selection processes among candidates for public research grants: the case of the Ramon y Cajal Programme in Spain. RESEARCH EVALUATION, 18 (2), 153-161
Cartlidge, E (2007) Peer review steps out of the shadows. PHYSICS WORLD, 20 (1), 29-30
Clerides, S; Pashardes, P; Polycarpou, A (2011) Peer Review vs Metric-based Assessment: Testing for Bias in the RAE Ratings of UK Economics Departments. ECONOMICA, 78, 565–583
Crystal, D (2003) English as a Global Language, Cambridge University Press.
Danish Agency for Science Technology and Innovation (2009) Research Evaluation: Methods, Practice, and Experience, Research: Analysis and Evaluation 1/2009, ed. by H. Foss Hansen.
Daryl E., Chubin, Hackett E. J. (1990). Peerless science: peer review and U.S. science policy. State University of New York Press: Albany. Retrieved March 4, 2012 from:
David, ME (2008) Research quality assessment and the metrication of the social sciences. EUROPEAN POLITICAL SCIENCE, 7 (1), 52-63
Davo, MD; Vives, C; Alvarez-Dardet, C (2003) Why are women underused in the JECH peer review process? JOURNAL OF EPIDEMIOLOGY AND COMMUNITY HEALTH, 57 (12), 936-937
Deibarrola, M (1992) Mexico - the experience of homologation and de-homologation on the wages for academic work. INTERCIENCIA, 17 (6), 348-353
Dolan, C. (2007) Feasibility Study: the Evaluation and Benchmarking of Humanities Research in Europe, Bristol: Arts & Humanities Research Council.
Eisenhart, M (2002) The paradox of peer review: admitting too much or allowing too little? RESEARCH IN SCIENCE EDUCATION, 32 (2), 241-255
European Science Foundation (2006) Peer Review. Its present and future state, Conference Report, Prague, 12-13 October 2006, Strasbourg.
ESF MO Forum (2009) Evaluation in National Research Funding Agencies: approaches, experiences and case studies. A report of the ESF Member Organisation Forum on Ex-Post Evaluation of Funding Schemes and Research Programmes. Strasbourg
ESF (2011). ESF Survey Analysis Report on Peer Review Practices. European Science Foundation: Strasbourg.
ESF Guide (2011). European Peer Review Guide Integrating Policies and Practices into Coherent Procedures. European Science Foundation: Strasbourg.
ESF MO Forum (2012) A Guide to improve Quality of ex-post Evaluation. Working Group 1: «Quality Assurance and Evaluation Guidelines» Member Organisation Forum on Publicly Funded Research. 27 June 2012
ESF MO Forum Report (2012) A report by the ESF Member Organisation Forum on Evaluation of Publicly Funded Research. 19 June 2012
ERC (2012) European Research Council. http://erc.europa.eu/ Visited 01 May 2012.
EUROSTAT (2009). Science, technology and innovation in Europe. Eurostat pocket books. Luxembourg: Office for Official Publications of the European Communities.
Fitzpatrick, K (2010) Peer-to-peer Review and the Future of Scholarly Authority Social Epistemology. A Journal of Knowledge, Culture and Policy, 24 (3), 161-179
Foss Hansen H. (2009). Research Evaluation: Methods, Practice, and Experience Research: Analysis and Evaluation. Danish Agency for Science, Technology and Innovation: Copenhagen.
Franks, AL; Simoes, EJ; Singh, R; Gray, BS (2006) Assessing prevention research impact - A bibliometric analysis. AMERICAN JOURNAL OF PEER REVIEWENTIVE MEDICINE, 30 (3), 211-216
Garcia-Aracil, A; Gracia, AG; Perez-Marin, M (2006) Analysis of the evaluation process of the research performance: An empirical case. SCIENTOMETRICS, 67 (2), 213-230

Georghiou, L. (1995).Research evaluation in European national science and technology systems. *RESEARCH EVALUATION,* 5(1*),* 3-10
Glaser, J (2004) Why are the most influential books in Australian sociology necessarily the most cited ones? JOURNAL OF SOCIOLOGY, 40 (3), 261-282
Haeffner-Cavaillon, N; Graillot-Gak, C (2009) The use of bibliometric indicators to help peer-review assessment. ARCHIVUM IMMUNOLOGIAE ET THERAPIAE EXPERIMENTALIS, 57 (1), 33-38
HCSRC - The Origin of the Scientific Journal and the Process of Peer Review, House of Commons Select Committee Report available at http://eprints.soton.ac.uk/263105/1/399we23.htm
Hegde, D (2009) Political Influence behind the Veil of Peer Review: An Analysis of Public Biomedical Research Funding in the United States. JOURNAL OF LAW & ECONOMICS, 52 (4), 665-690
Herzog, HA; Podberscek, AL; Docherty, A (2005) The reliability of peer review in anthrozoology. ANTHROZOOS, 18 (2), 175-182
Hicks, D; Tomizawa, H; Saitoh, Y; Kobayashi, S (2004) Bibliometric techniques in the evaluation of federally funded research in the United States. RESEARCH EVALUATION, 13 (2), 78-86
Hodgson, C (1995) Evaluation Of Cardiovascular Grant-In-Aid Applications By Peer-Review - Influence Of Internal And External Reviewers And Committee. SCANADIAN JOURNAL OF CARDIOLOGY, 11 (10), 864-868
Hojat, M; Gonnella, JS; Caelleigh, AS (2003) Impartial judgment by the "gatekeepers" of science: Fallibility and accountability in the peer review process. ADVANCES IN HEALTH SCIENCES EDUCATION, 8 (1), 75-96
Holbrook, A (2000) Evaluation of research sponsored by federal granting councils in Canada: the social contract. RESEARCH EVALUATION, 9 (1), 47-56
Horton, R (2000) Genetically modified food: consternation, confusion, and crack-up. *MJA* 172 (4), 148–9
http://books.google.ee/books?id=Xfsh6D29WoIC&lpg=PP1&dq=%22Peer+review%22&pg=PP1&redir_esc=y#v=onepage&q&f=true
Ismail,Sh; Farrands, A; Wooding, S (2009) Evaluating Grant Peer Review in the Health Sciences A review of the literatuure. RAND Report TR-742-DH
Jansen, D; Heidler, R; von Goertz, R (2009) Unequal chances in the science system: artefact or reality? KOLNER ZEITSCHRIFT FUR SOZIOLOGIE UND SOZIALPSYCHOLOGIE, 61 (3), 463-467
Jayasinghe, UW; Marsh, HW; Bond, N (2006) A new reader trial approach to peer review in funding research grants: An Australian experiment. SCIENTOMETRICS, 69 (3), 591-606
Kaplan, D; Lacetera, N; Kaplan, C (2008) Sample Size and Precision in NIH Peer Review. PLOS ONE, 3 (7)
Kenna, R; Berche, B (2011) Normalization of peer-evaluation measures of group research quality across academic disciplines. RESEARCH EVALUATION, 20 (2), 107-116
Langfeldt, L (2001) The decision-making constraints and processes of grant peer review, and their effects on the review outcome. SOCIAL STUDIES OF SCIENCE, 31 (6), 820-841
Laudel, G (2006) Conclave in the Tower of Babel: how peers review interdisciplinary research proposals. RESEARCH EVALUATION, 15 (1), 57-68
Laudel, G (2006) The 'quality myth': Promoting and hindering conditions for acquiring research funds. HIGHER EDUCATION, 52 (3), 375-403
Lee, M; Om, K; Koh, J (2000) The bias of sighted reviewers in research proposal evaluation: A comparative analysis of blind and open review in Korea. SCIENTOMETRICS, 48 (1), 99-116
Levitt, JM; Thelwall, M (2011) A combined bibliometric indicator to predict article impact. INFORMATION PROCESSING AND MANAGEMENT, 47, 300-308
Lincoln, A; Pincus, S; Koster, J; Leboy, P (2012) The Matilda Effect in science: Awards and prizes in the US, 1990s and 2000s. SOCIAL STUDIES OF SCIENCE, 42(2), 307-320
Lovegrove, BG; Johnson, SD (2008) Assessment of research performance in biology: How well do peer review and bibliometry correlate? BIOSCIENCE, 58 (2), 160-164
Lundberg, J (2007) Lifting the crown-citation z-score. JOURNAL OF INFORMETRICS, 1 (2), 145-154
Luukkonen, T (2002) Research evaluation in Europe: state of the art. RESEARCH EVALUATION, 11 (2), 81-84

Mansilla, VB (2006) Assessing expert interdisciplinary work at the frontier: an empirical exploration. RESEARCH EVALUATION, 15 (1), 17-29
Mackiewcz, W; Hazelkorn, E; Berghoff, S; Bonaccorsi, A; Borrell-Damian, L; Emplit, Ph; Inzelt, A; Marklund, G; Moed, H; Mustajokl, A; Newby, H; Peyraube, A; Rowley, S; Salmi, J; Sanchez, P (2010) Assessing Europe's University-Based Research Expert Group on Assessment Luxembourg: Publications Office of the European Union.
Marsh, HW; Jayasinghe, UW; Bond, NW (2008) Improving the peer-review process for grant applications - Reliability, validity, bias, and generalizability. AMERICAN PSYCHOLOGIST, 63 (3), 160-168
Marsha, H; Jayasinghe, UW; Bond, NW (2011) Gender differences in peer reviews of grant applications: A substantive-methodological synergy in support of the null hypothesis model. JOURNAL OF INFORMETRICS, 5, 167-180
Meho, LI; Sonnenwald, DH (2000) Citation ranking versus peer evaluation of senior faculty research performance: A case study of Kurdish scholarship. JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE, 51 (2), 123-138.
Merton, R (1968). The Matthew Effect in Science. Science, 159 (3810), 56–63
Merton, R (1988) The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. ISIS, 79, 606-623
Molinie, A; Bodenhausen, G (2010) Bibliometrics as Weapons of Mass Citation. CHIMIA, 64, 78-89
Muench, R (2010) The monopoly mechanism in science. On the shoulders of Robert K. Merton. BERLINER JOURNAL FUR SOZIOLOGIE, 20 (3), 341-370
NordForsk NORIA-net (2010) Development of Peer Review in the Nordic Context. Report. Helsinki. http://193.167.96.163/Tiedostot/Tiedostot/Julkaisut/Development_of_Peer_Review_in_the_Nordic_Context.pdf
NORIA-net (2010). Development of Peer Review in the Nordic Context REPORT. NordForsk NORIA-net: Helsinki.
Norris, M; Oppenheim, C (2010) Peer review and the h-index: Two studies. JOURNAL OF INFORMETRICS, 4 (3), 221-232
Olsson, G (1999) Women in the university - guests at the men's table. NORDISK PSYKOLOGI, 51 (1), 59-76
Oswald, AJ (2010) A suggested method for the measurement of world-leading research (illustrated with data on economics). SCIENTOMETRICS, 84 (1), 99-113
Oxman, AD; Guyatt, GH (1991) Validation Of An Index Of The Quality Of Review Articles. JOURNAL OF CLINICAL EPIDEMIOLOGY, 44 (11), 1271-1278
Pendlebury, DA (2009) The use and misuse of journal metrics and other citation indicators. ARCHIVUM IMMUNOLOGIAE ET THERAPIAE EXPERIMENTALIS, 57 (1), 1-11

PROTOCOL FOR RESEARCH ASSESSMENT IN THE NETHERLANDS Published VSNU, KNAW and NWO 2009 (Updated June 2010) This document is available through internet only: www.knaw.nl/sep

Rasmussen, Jesper; Langer, Vibeke; Alroe, Hugo Fjelsted (2006) Bias in peer review of organic farming grant applications. AGRICULTURE AND HUMAN VALUES, 23 (2), 181-188

Refsdal, M. (2010) Peer review at the University of Tromsø : a study of time spent on reviewing and researchers' opinions on peer review. Master thesis. http://hdl.handle.net/10037/2760

Reinhart, M. (2010). Peer review practices: a content analysis of external reviews in science funding. *RESEARCH EVALUATION,* 19(5), 317-331
Retzer, V.; Jurasinski, G. (2009). Towards objectivity in research evaluation using bibliometric indicators – A protocol for incorporating complexity. BASIC AND APPLIED ECOLOGY, 10(5), 393-400
Richards, K; Batty, M; Edwards, K; Findlay, A; Foody, G; Frostick, L; Jones, K; Lee, R; Livingstone, D; Marsden, T; Petts, J; Philo, C; Simon, D; Smith, S; Thomas, D (2009) The nature of publishing and assessment in Geography and Environmental Studies: evidence from the Research Assessment Exercise 2008. AREA, 41 (3), 231-243
Rinia, EJ; van Leeuwen, TN; van Vuren, HG; van Raan, AFJ (1998) Comparative analysis of a set of bibliometric indicators and central peer review criteria - Evaluation of condensed matter physics in the Netherlands. RESEARCH POLICY, 27 (1), 95-107
Rinia, EJ; van Leeuwen, TN; van Vuren, HG; van Raan, AFJ (2001) Influence of interdisciplinarity on peer-review and bibliometric evaluations in physics research. RESEARCH POLICY, 30 (3), 357-361

Rinia, EJ; van Leeuwen, TN; van Vuren, HG; van Raan, AFJ (2001) Influence of interdisciplinarity on peer-review and bibliometric evaluations in physics research. RESEARCH POLICY, 30 (3), 357-361
Sandstrom, U; Hallsten, M (2008) Persistent nepotism in peer-review. SCIENTOMETRICS, 74 (2), 175-189
Scantlebury, K (2002) A snake in the nest or in a snake's nest: What counts as peer review for a female science educator in a chemistry department? RESEARCH IN SCIENCE EDUCATION, 32 (2), 157-162
Scott, A (2007) Peer review and the relevance of science. FUTURES, 39 (7), 827-845
She Figures (2009). *She Figures 2009.* Statistics and Indicators on Gender Equality in Science. Luxembourg: Publications Office of the European Union.
Singapore Statement on Research Integrity http://www.singaporestatement.org/
SKEP ERA-NET Work Package 3 (2008) Approaches and practices in the evaluation of environmental research programmes. THE FINNISH ENVIRONMENT 13
Slesser, A; Qureshi, Y (2009) The Implications of Fraud in Medical and Scientific Research . WORLD JOURNAL OF SURGERY, 33 (11), 2355-2359
So, CYK (1998) Citation ranking versus expert judgment in evaluating communication scholars: Effects of research specialty size and individual prominence. SCIENTOMETRICS, 41 (3), 325-333
Souder, L (2011) The ethics of scholarly peer review: a review of the literature. LEARNED PUBLISHING, 24, 55–74
Spier, R. (2002). The history of the peer-review process. TRENDS IN BIOTECHNOLOGY, 20 (8), 357-358
Squazzoni, F; Gandelli, C (2012) Saint Matthew strikes again: An agent-based model of peer review and the scientific community structure. JOURNAL OF INFORMETRICS, 6 (2), 265-275
Taylor, J (2011) The Assessment of Research Quality in UK Universities: Peer Review or Metrics? BRITISH JOURNAL OF MANAGEMENT, 22, 202-217
Tite, L.; Schroter, S. (2007) Why do peer reviewers decline to review? A survey. JOURNAL OF EPIDEMIOLOGY AND COMMUNITY HEALTH, 61(1), 9-12.
Van Looy, B; Ranga, M; Callaert, J; Debackere, K; Zimmermann, E (2004) Combining entrepreneurial and scientific performance in academia: towards a compounded and reciprocal Matthew-effect? RESEARCH POLICY, 33 (3), 425-441
van Raan, AFJ (1996) Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. SCIENTOMETRICS, 36 (3), 397-420
van Raan, AFJ (2000) The pandora's box of citation analysis: Measuring scientific excellence - The last evil? ASIST MONOGRAPH SERIES, 301-319
van Raan, AFJ (2006) Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. SCIENTOMETRICS, 67 (3), 491-502
van Rooyen, S (2001) The evaluation of peer-review quality. LEARNED PUBLISHING, 14 (2), 85-91
van Rooyen,S; Godlee,F; Evans,S; Black, N; Smith, R (1999) Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial. BRITISH MEDICAL JOURNAL, 318 (7175), 23-27
Wenneras, C; Wold, A (1997) Nepotism and sexism in peer-review. NATURE, 387, 341-343
Wu, H; Ismail,S; Guthrie, S; Wooding, S (2011) Alternatives to Peer Review in Research Project Funding. RAND, pp 28.

Annex 1

ACUMEN Peer Review Questionnaire Survey for Funding Grants

Peer review is a natural double faceted part of the researcher's working cycle. One facet allows the researcher to be involved in the process as the author (co-author) of papers submitted to conferences and journals, and when applying for academic positions and obtaining research grants/contracts and awards. The other facet enables the researcher to participate in the process as a reviewer, editor or a funder.

In this questionnaire, we are interested in your views and experiences on Peer review as a project/grant applicant or reviewer.

Answering the survey will take approximately 20 minutes.

There are 55 questions in this survey

1. General information

1.1. Your fields of science and technology

Please choose only one of the following:

- ○ Natural Sciences
- ○ Engineering and Technology
- ○ Medical Sciences
- ○ Agricultural Sciences
- ○ Social Sciences
- ○ Humanities

1.1.6. Humanities

Only answer this question if the following conditions are met:
° Answer was 'Humanities' at question '1 [q.1.1]' (1.1. Your fields of science and technology)

Please choose only one of the following:

- ○ History
- ○ Languages and literature
- ○ Other humanities

1.1.5. Social sciences

Only answer this question if the following conditions are met:
° Answer was 'Social Sciences' at question '1 [q.1.1]' (1.1. Your fields of science and technology)

Please choose only one of the following:

- ○ Psychology
- ○ Economics
- ○ Educational sciences
- ○ Other social sciences

1.1.4. Agricultural sciences

Only answer this question if the following conditions are met:
° Answer was 'Agricultural Sciences' at question '1 [q.1.1]' (1.1. Your fields of science and technology)

Please choose only one of the following:

- ○Agriculture, forestry, fisheries and allied sciences
- ○Veterinary medicine

1.1.3. Medical sciences

Only answer this question if the following conditions are met:
° Answer was 'Medical Sciences' at question '1 [q.1.1]' (1.1. Your fields of science and technology)

Please choose only one of the following:

- ○Basic medicine
- ○Clinical medicine
- ○Health sciences

1.1.1 Natural sciences

Only answer this question if the following conditions are met:
° Answer was 'Natural Sciences' at question '1 [q.1.1]' (1.1. Your fields of science and technology)

Please choose only one of the following:

- ○Mathematics and computer sciences
- ○Physical sciences
- ○Chemical sciences
- ○Earth and related environmental sciences
- ○Biological sciences

1.1.2 Engineering and technology

Only answer this question if the following conditions are met:
° Answer was 'Engineering and Technology' at question '1 [q.1.1]' (1.1. Your fields of science and technology)

Please choose only one of the following:

- ○Civil engineering
- ○Electrical engineering, electronics
- ○Other engineering sciences

1.2. Country of Affiliation (present)

Please choose only one of the following:

- ○ Afghanistan
- ○ Albania
- ○ Algeria
- …

1.3. Citizenship

Please choose only one of the following:

- ○ Afghanistan
- ○ Albania
- ○ Algeria
- …

1.4. Gender

Please choose only one of the following:

- ○ Female
- ○ Male

1.5. What is your academic position?

Please choose all that apply:

- ☐ Full professor
- ☐ Associate professor / Reader / Senior lecturer
- ☐ Lecturer / Assistant professor
- ☐ Postdoctoral research fellow
- ☐ Student (e.g. PhD. or Master)
- ☐ Other:

2. Participation in the Peer Review process

2.1. Have you been in the following roles: *

Please choose the appropriate response for each item:

	Yes	No
Applicant	○	○
Reviewer	○	○

3. Your experience as an applicant or author in the Peer Review process

3.1. Please recall how many applications you have made during your scientific career.

Personal Grant or Fellowship

Please choose the appropriate response for each item:

	None	Less than 10	Dozens	Hundred(s)
National	○	○	○	○
International	○	○	○	○
Supranational (EU FP, ESF)	○	○	○	○

3.1.2 Please recall how many applications have you made during your scientific career.

As coordinator of a project

Please choose the appropriate response for each item:

	None	Less than 10	Dozens	Hundred(s)
National	○	○	○	○
International	○	○	○	○

	None	Less than 10	Dozens	Hundred(s)
Supranational (EU FP, ESF)	○	○	○	○

3.2.1. Please estimate (in %, approximately) how successful on an average have your applications been in the last five years.

Please write a multiple of 5 between 0 and 100 for each item:

success % of National applications

Personal Grant of Fellowship

Coordinator of a collaborative project

3.2.2. Please estimate (in %, approximately) how successful on an average have your applications been in the last five years.

Please write a multiple of 5 between 0 and 100 for each item:

success % of International applications

Personal Grant of Fellowship

Coordinator of a collaborative project

3.2.3. Please estimate (in %, approximately) how successful on an average have your applications been in the last five years.

Please write a multiple of 5 between 0 and 100 for each item:

Success % of Supranational applications

Personal Grant of Fellowship

Coordinator of a collaborative project

3.3. Were you informed of the reasons for rejection?

Please choose only one of the following:

- ○ Yes, always
- ○ Yes, in most cases
- ○ Yes, but in a few cases
- ○ No

3.4.1. What were the reasons for the rejection of your (national) applications (last five years)?

1 = very rarely, 2 = rarely, 3 = often, 4 = very often

Please choose the appropriate response for each item:

	1	2	3	4
Did not pass eligibility check	○	○	○	○
Did not pass evaluation treshold	○	○	○	○
Did not pass the interview	○	○	○	○
Lack of funding	○	○	○	○

3.4.2. What were the reasons for the rejection of your (international) applications (last five years)?

1 = very rarely, 2 = rarely, 3 = often, 4 = very often

Please choose the appropriate response for each item:

	1	2	3	4
Did not pass eligibility check	○	○	○	○
Did not pass evaluation treshold	○	○	○	○
Did not pass the interview	○	○	○	○
Lack of funding	○	○	○	○

3.4.3. What were the reasons for the rejection of your (supranational) applications (last five years)?

1 = very rarely, 2 = rarely, 3 = often, 4 = very often

Please choose the appropriate response for each item:

	1	2	3	4
Did not pass eligibility check	○	○	○	○
Did not pass evaluation treshold	○	○	○	○
Did not pass the interview	○	○	○	○
Lack of funding	○	○	○	○

3.5. What is your best experience of Peer Review? Could it be used as a model for good practice? If yes, please describe it briefly (*For example: Academy of Finland, Academy Professors funding instrument, ...*)

Please write your answer here:

3.6. What problems have you encountered as an applicant? Do you have suggestions for improving the system?

Please write your answer here:

4. Your experience in the Peer Review process as a reviewer

4.1. Please recall how many reviews you have made during your scientific career.

Allocation of research funding (grants and projects)

Please choose the appropriate response for each item:

	None	Less than 10	Dozens	Hundred(s)
National	○	○	○	○

	None	Less than 10	Dozens	Hundred(s)
International	○	○	○	○
Supranational (EU FP)	○	○	○	○

4.2.1. What were the main reasons for consenting to be a reviewer (last five years)?

National reviews

1 = very rarely, 2 = rarely, 3 = often, 4 = very often

Please choose the appropriate response for each item:

	1	2	3	4
I feel a general obligation towards the field	○	○	○	○
It gives me an overview of my field	○	○	○	○
It improves my reputation	○	○	○	○
It allows me to stop or hinder my competition	○	○	○	○
I want to ensure the quality of my field	○	○	○	○
I want to educate/help fellow researchers	○	○	○	○
It is cost effective	○	○	○	○

4.2.2. What were the main reasons for consenting to be a reviewer (last five years)?

International reviews

1 = very rarely, 2 = rarely, 3 = often, 4 = very often

Please choose the appropriate response for each item:

	1	2	3	4
I feel a general obligation towards the field	○	○	○	○
It gives me an overview of my field	○	○	○	○
It improves my reputation	○	○	○	○
It allows me to stop or hinder my competition	○	○	○	○
I want to ensure the quality of my field	○	○	○	○
I want to educate/help fellow researchers	○	○	○	○
It is cost effective	○	○	○	○

4.2.3. What were the main reasons for consenting to be a reviewer (last five years)?

Supranational reviews

1 = very rarely, 2 = rarely, 3 = often, 4 = very often

Please choose the appropriate response for each item:

	1	2	3	4
I feel a general obligation towards the field	○	○	○	○
It gives me an overview of my field	○	○	○	○
It improves my reputation	○	○	○	○
It allows me to stop or hinder my competition	○	○	○	○
I want to ensure the quality of my field	○	○	○	○

	1	2	3	4
I want to educate/help fellow researchers	○	○	○	○
It is cost effective	○	○	○	○

4.3. Have you refused a request to review an application (in the past five years)?

Please choose only one of the following:

- ○ Yes, often
- ○ Yes, rarely
- ○ Never

4.4.1. What were the main reasons for refusal (last five years)?

National reviews

1 = very rarely, 2 = rarely, 3 = often, 4 = very often

Please choose the appropriate response for each item:

	1	2	3	4
I didn't have time due to other commitments	○	○	○	○
I did not have relevant expertise	○	○	○	○
I found it boring	○	○	○	○
There was a personal conflict of interest between me and the applicant	○	○	○	○
There was an institutional conflict of interest	○	○	○	○
My institution's policy on time allocation means I could not justify the time	○	○	○	○
No payment was offered (for me	○	○	○	○

	1	2	3	4
personally)				
No payment was offered (for my institution)	○	○	○	○

4.4.2. What were the main reasons for refusal (last five years)?

International reviews

1 = very rarely, 2 = rarely, 3 = often, 4 = very often

Please choose the appropriate response for each item:

	1	2	3	4
I didn't have time due to other commitments	○	○	○	○
I did not have relevant expertise	○	○	○	○
I found it boring	○	○	○	○
There was a personal conflict of interest between me and the applicant	○	○	○	○
There was an institutional conflict of interest	○	○	○	○
My institution's policy on time allocation means I could not justify the time	○	○	○	○
No payment was offered (for me personally)	○	○	○	○
No payment was offered (for my institution)	○	○	○	○

4.4.3. What were the main reasons for refusal (last five years)?

Supranational reviews

1 = very rarely, 2 = rarely, 3 = often, 4 = very often

Please choose the appropriate response for each item:

	1	2	3	4
I didn't have time due to other commitments	○	○	○	○
I did not have relevant expertise	○	○	○	○
I found it boring	○	○	○	○
There was a personal conflict of interest between me and the applicant	○	○	○	○
There was an institutional conflict of interest	○	○	○	○
My institution's policy on time allocation means I could not justify the time	○	○	○	○
No payment was offered (for me personally)	○	○	○	○
No payment was offered (for my institution)	○	○	○	○

4.5. Have you been informed of the final results of the application reviewed by you (last five years)?

Please choose the appropriate response for each item:

	National	International	Supranational
Yes, always	○	○	○
Yes, often	○	○	○
Yes, rarely	○	○	○
Never	○	○	○

4.6. Would this feedback about the final results be necessary?

Please choose all that apply and provide a comment:

- ☐ Yes (please explain)
- ☐ No (please explain)

4.7. Do you have any additional comments on this issue?

Please write your answer here:

5. Most appropriate indicators

Some research assessments assign different weightings or values to the various indicators. In this way, some components of research activity are valued more highly than other activities.

5.1. Please rate (from 1 to 5) what are the most appropriate indicators in assessing the work of reasearchers.

1 = lowest rating, 5 = highest rating

5.1.1. Input indicators

Please choose the appropriate response for each item:

	1	2	3	4	5
External funding	○	○	○	○	○
Recruitment of PhD students	○	○	○	○	○

5.1.2. Process indicators

Please choose the appropriate response for each item:

	1	2	3	4	5
Seminar and conference activity	○	○	○	○	○

	1	2	3	4	5
Invited keynotes	○	○	○	○	○
International visiting research appointments	○	○	○	○	○

5.1.3. Structure indicators

Please choose the appropriate response for each item:

	1	2	3	4	5
Number of PhD students	○	○	○	○	○
Research collaborations and partnership	○	○	○	○	○
Reputation and esteem (position as journal editors, membership of editorial boards and scientific committees and membership in learned academies)	○	○	○	○	○

5.1.4. Output indicators

Please choose the appropriate response for each item:

	1	2	3	4	5
Publications	○	○	○	○	○
Non-bibliographical outputs	○	○	○	○	○
Number of PhD graduates and completion rates for graduates	○	○	○	○	○
Patent development	○	○	○	○	○
Public outreach (broader dissemination, transfer and exchange of research results)	○	○	○	○	○

5.1.5. Effect indicators

Please choose the appropriate response for each item:

	1	2	3	4	5
Citations	○	○	○	○	○

	1	2	3	4	5
H index	○	○	○	○	○
Number of prestigious national and international awards and prizes	○	○	○	○	○
Employability of PhD graduates (in private sector)	○	○	○	○	○

5.1.6. Social indicators

Please choose the appropriate response for each item:

	1	2	3	4	5
Relevance to citizens' concerns	○	○	○	○	○
Relevance to global societal challenges	○	○	○	○	○
Usefulness to policy decision makers	○	○	○	○	○
Relevance to science communication initiatives	○	○	○	○	○
Contributing to science education	○	○	○	○	○

5.2. Should the various indicators be weighed differently?

Please choose all that apply and provide a comment:

- ☐ Yes (please explain)
- ☐ No

5.3. Is there a need to have different weightings for the various indicators in different subject areas?

Please choose all that apply and provide a comment:

- ☐ Yes (please explain)
- ☐ No

5.4. Is there a need to have various indicators or different weighing for the various indictors at different career stages?

Please choose all that apply and provide a comment:

- ☐ Yes (please explain)
- ☐ No

5.5. Do you have any additional comments on this issue?

Please write your answer here:

6. Problems with Peer Review

Peer review relies on mutual trust and honesty, hundreds of articles and books have been written about peer review (https://www.archimedes.ee/acumen). Although generally considered essential to academic quality and used widely, peer review has also been criticized.

6.1. Please rate (from 1 to 5) the extent to which the following bias may affect the assessment of the applications

(1 = no problem, 5 = requires major changes)

Please choose the appropriate response for each item:

	1	2	3	4	5
Institutional bonus (scientists from prestigious institutions)	○	○	○	○	○
Geographical origin	○	○	○	○	○
Language (favouring of papers written in English)	○	○	○	○	○
Conflict of interest	○	○	○	○	○
Gender	○	○	○	○	○
The scope of the research	○	○	○	○	○
Peer Review as a time consuming process	○	○	○	○	○
High cost of Peer Review	○	○	○	○	○
So-called Matthew effect: "to those who have, more shall be given"	○	○	○	○	○
So-called friendship bonus	○	○	○	○	○
Entrenched academic traditionalism	○	○	○	○	○

6.2. What are the most important issues to consider for optimizing fairness and objectivity in the evaluation process?

Please write your answer here:

6.3. What are the most essential criteria for a good reviewer?

Please rate from 1 to 5 (1 = not important, 5 = very important)

Please choose the appropriate response for each item:

	1	2	3	4	5
Academic excellence	○	○	○	○	○
Previous peer review experience	○	○	○	○	○
Relevant disciplinary (subject area) competence	○	○	○	○	○
Timeliness: was the assessment sent in time	○	○	○	○	○
Comprehensibility of the comments provided: was it possible to understand the comments?	○	○	○	○	○
Usefulness: was the assessment fit for purpose?	○	○	○	○	○
Appropriateness of the language used: i.e., was the assessment free of disrespectful offensive comments towards the proposers?	○	○	○	○	○

6.4. What do you think to what extent information about reviewers should be available to the applicants?

Please choose all that apply:

- ☐ A reviewer's written evaluation is available to the applicant, including reviewer's name
- ☐ A reviewer's written evaluation is available to the applicant, excluding reviewer's name
- ☐ A reviewer's rating (grade) is available to the applicant
- ☐ Right to reply - applicant has the possibility to read and respond to the reviewer's comment before the final decision
- ☐ A reviewer is anonymous to the applicant, however, list of reviewers is published after call's financing decisions are made

6.5. Have you had experience of a system that allows applicants to nominate possible reviewers?

Please choose only one of the following:

- ○ Yes
- ○ No

6.6. If yes, did you nominate reviewers?

Please choose all that apply and provide a comment:

- ☐ Yes (please specify)
- ☐ No (please specify)

6.7. Have you had experience of a system that allows applicants to exclude reviewers?

Please choose only one of the following:

- ○ Yes
- ○ No

6.8. If yes, did you give names of potential reviewers to be excluded?

Please choose all that apply and provide a comment:

- ☐ Yes (please specify)
- ☐ No (please specify)

6.9. What is your view of such systems?

Please write your answer here:

7. Scholarly Peer Review has been under critical attention of researchers from its beginning. There

7.1. What is your opinion about the future of the peer review system?

Please choose all that apply and provide a comment:

- ☐ Current system is sufficient
- ☐ There is a need for improvements (please specify)

7.2. Do you have any further comments on peer review?

Please write your answer here:

Thank You for answering!

Survey results of the analysis will be published at http://research-acumen.eu in May 2012

Annex 2. Questionnaire responses

1.1. Your fields of science and technology

Field	Count
Agricultural Sciences	58
Engineering and Technology	287
Humanities	118
Medical Sciences	234
Natural Sciences	1132
Social Sciences	199

1.2 and 1.3. Country of affiliation and citizenship

Country	Affiliation	Citizenship
UK	272	120
Spain	223	244
Germany	208	201
France	190	176
Italy	182	261
Netherlands	103	89
Switzerland	83	28
Israel	77	79
USA	58	55
Belgium	55	51
Sweden	54	34

Poland	43	68
Greece	40	70
Denmark	39	28
Finland	33	25
Portugal	33	35
Austria	32	24
Turkey	28	30
Ireland	26	20
Hungary	22	27
Norway	22	17
Australia	16	14
China	15	14
Bulgaria	13	21
Czech Republic	13	14
Russia	13	28
Ukraine	12	23
Canada	11	17
Japan	10	12
Cyprus	9	9
Estonia	9	9
Romania	9	18
India	8	22
n/a	8	0
Lithuania	7	8
Mexico	6	8

Slovakia	6	8
Slovenia	5	9
Argentina	4	8
Brazil	4	6
Georgia	4	5
Egypt	3	3
Iceland	3	3
New Zealand	3	3
Tunisia	3	3
Chile	2	3
Croatia	2	2
Nigeria	2	2
Qatar	2	0
Uruguay	2	4
Algeria	1	2
Andorra	1	1
Belarus	1	2
Cuba	1	2
Indonesia	1	1
Iran	1	3
Malaysia	1	1
Papua New Guinea	1	3
Philippines	1	1
Republic of Korea	1	1
Serbia	1	4

Singapore	1	2
Thailand	1	1
The former Yugoslav Republic of Macedonia	1	2
Trinidad and Tobago	1	1
Uzbekistan	1	2
Vietnam	1	2
Albania	0	4
Bangladesh	0	1
Bosnia and Herzegovina	0	1
Colombia	0	1
Ecuador	0	1
Guatemala	0	1
Latvia	0	3
Luxembourg	0	1
Madagascar	0	1
Malta	0	2
Morocco	0	1
Mozambique	0	1
Peru	0	4
Venezuela	0	1
Total	2044	2012

1.4. Gender

Gender	Count

Female	652
Male	1337
n/a	125

1.5. What is your academic position?

Academic position	Count
Full professor	439
Associate professor / Reader / Senior lecturer	335
Lecturer / Assistant professor	338
Postdoctoral research fellow	668
Student (e.g. PhD. or Master)	30

2.1. Have you been in the following roles: Applicant and Reviewer

Role	Applicant	Reviewer
Yes	1975	889
No	76	1161

3.1. Please recall how many applications you have made during your scientific career

	None	Less than 10	Dozens	Hundreds
Personal Grant or Fellowship [National]	226	1051	400	14
Personal Grant or Fellowship [International]	518	1002	169	3
Personal Grant or Fellowship [Supranational (EU FP, ESF)]	283	1301	100	1
As coordinator of a project [National]	623	816	244	3
As coordinator of project	983	650	55	0

[International]				
As coordinator of project [Supranational (EU FP, ESF)]	910	742	28	0

3.2. Please estimate (in %, approximately) how successful on an average have your applications been

	0	5–20	25–45	50–70	75–95	100
Personal Grant or Fellowship [National]	145	126	171	376	272	266
Personal Grant or Fellowship [International]	164	114	119	261	135	205
Personal Grant or Fellowship [Supranational (EU FP, ESF)]	152	78	123	284	105	437
As coordinator of a project [National]	216	107	126	242	152	158
As coordinator of project [International]	199	87	73	158	70	107
As coordinator of project [Supranational (EU FP, ESF)]	227	78	75	156	58	148

3.3. Were you informed of the reasons for rejection?

	Count
No	119
Yes, always	439
Yes, but in a few cases	279
Yes, in most cases	631

4.1. Please recall how many reviews you have made during your scientific career. Allocation of research funding (grants and projects)

Type	None	Less than 10	Dozens	Hundreds
Grants and Projects [National]	164	264	198	57
Grants and Projects [International]	150	338	173	23
Grants and Projects [Supranational (EU FP, ESF)]	375	210	89	9

3.4. What were the reasons for the rejection of your applications (last five years)? 1 = very rarely, 2 = rarely, 3 = often, 4 = very often

	National				International				Supranational			
	Very rarely	Rarely	Often	Very often	Very rarely	Rarely	Often	Very often	Very rarely	Rarely	Often	Very often
Did not pass eligibility check	599	40	31	30	425	41	29	15	392	29	22	15
Did not pass evaluation threshold	268	205	239	204	198	159	205	149	199	148	182	150
Did not pass the interview	340	88	57	31	261	61	56	22	243	45	44	20
Lack of funding	174	100	213	324	160	77	131	170	136	68	110	189

4.2. What were the main reasons for consenting to be a reviewer (last five years)? 1 = very rarely, 2 = rarely, 3 = often, 4 = very often

	National				Inter-national				Supra-national			
	1	2	3	4	1	2	3	4	1	2	3	4
I feel a general obligation towards the field	9	22	127	339	12	23	132	334	7	15	67	193
It gives me an overview of my field	55	113	155	153	46	97	178	162	24	49	78	112
It improves my reputation	147	138	98	61	141	118	118	74	75	78	49	43
It allows me to stop or hinder my competition	326	29	12	1	332	28	12	6	172	18	7	6
I want to ensure the quality of my field	49	68	182	168	52	79	189	145	31	48	93	90
I want to educate/help fellow researchers	81	113	164	87	90	122	152	90	60	60	79	54
It is cost effective	251	62	39	17	242	72	35	17	110	37	33	22

4.3. Have you refused a request to review an application (in the past five years)?

Answer	Count
Never	273
Yes, rarely	330
Yes, often	82

4.4. What were the main reasons for refusal (last five years)? 1 = very rarely, 2 = rarely, 3 = often, 4 = very often

	National				Inter+ national				Supra+ national			
	1	2	3	4	1	2	3	4	1	2	3	4
I didn't have time due to other commitments	26	26	51	226	27	22	47	217	19	14	29	129
I did not have relevant expertise	35	74	120	77	29	46	74	45	31	50	63	31
I found it boring	170	35	21	9	101	29	13	5	108	16	6	4
There was a personal conflict of interest between me and the applicant	110	88	46	38	88	45	18	15	79	44	18	23
There was an institutional conflict of interest	177	49	16	15	113	29	4	6	104	25	9	12
My institution's policy on time allocation (means I could not justify the time)	197	12	6	10	125	8	2	6	121	5	0	5
No payment was offered (for my institution)	205	14	6	2	129	8	2	1	123	7	1	1

4.5. Have you been informed of the final results of the application reviewed by you (last five years)?

	Never	Yes, rarely	Yes, often	Yes, always
National	170	103	124	110
International	163	138	128	82
Supranational	106	73	71	62

4.6. Would this feedback about the final results be necessary?

No	1593
Yes	522

5.1. Please rate what are the most appropriate indicators in assessing the work of researchers (1 = lowest rating, 5 = highest rating)

	1	2	3	4	5
Non-bibliographical outputs	146	351	486	265	74
Usefulness to policy decision makers	219	304	419	310	125
Employability of PhD graduates (in private sector)	251	276	429	300	124
Relevance to citizens' concerns	179	309	419	336	152
Patent development	278	274	333	296	176
Number of PhD students	142	327	580	357	74
Recruitment of PhD students	157	287	553	362	105
Relevance to science communication initiatives	139	229	414	452	156
Relevance to global societal challenges	142	261	380	398	210

Public outreach	115	272	444	426	204
Number of PhD graduates and completion rates for graduates	92	245	524	472	147
Seminar and conference activity	97	231	508	479	178
International visiting research appointments	91	250	484	460	193
External funding	86	184	443	484	286
Contributing to science education	87	162	370	487	310
Invited keynotes	80	202	418	516	263
Reputation and esteem	61	194	408	506	325
H index	104	141	353	425	296
Number of prestigious national and international awards and prizes	88	156	360	481	388
Research collaborations and partnership	31	132	358	633	343
Citations	55	94	317	494	520
Publications	15	21	114	411	943

5.2. – 5.4. Should the various indicators be weighed differently?

	Yes	No
1.2. Should the various indicators be weighed differently?	874	442
1.3. Is there a need to have different weightings for the various indicators in different subject areas?	912	418
1.4. Is there a need to have various indicators or different weighing for the various indictors at different career stages?	927	418

6.1. Please rate (from 1 to 5) the extent to which the following bias may affect the assessment of the applications (1 = no problem, 5 = requires major changes)

	1	2	3	4	5
Institutional bonus (scientists from prestigious institutions)	92	158	287	481	379
Geographical origin	203	246	385	334	224
Language (favouring of papers written in English)	289	182	257	361	301
Conflict of interest	166	268	378	290	253
Gender	645	297	250	116	65
The scope of the research	193	223	398	340	177
Peer Review as a time consuming process	162	236	421	341	187
High cost of Peer Review	288	306	349	228	93
Matthew effect	89	115	259	420	484
So-called friendship bonus	92	146	301	429	380
Entrenched academic traditionalism	98	127	363	404	313

6.3. What are the most essential criteria for a good reviewer? Please rate from 1 to 5 (1 = not important, 5 = very important)

	1	2	3	4	5
Academic excellence	20	44	179	523	678
Previous peer review experience	72	191	423	481	267
Relevant disciplinary (subject area) competence	17	16	96	405	915
Timeliness: was the assessment sent in time	32	137	426	552	280

Comprehensibility of the comments provided: was it possible to understand the comments?	11	26	148	578	672
Usefulness: was the assessment fit for purpose	8	29	170	545	643
Appropriateness of the language used: i.e., was the assessment free of disrespectful offensive comments towards the proposers	39	128	273	478	495

6.4. What do you think to what extent information about reviewers should be available to the applicants?

	No	Yes
A reviewer's written evaluation is available to the applicant, including reviewer's name	1895	219
A reviewer's written evaluation is available to the applicant, excluding reviewer's name	959	1155
A reviewer's rating (grade) is available to the applicant	1320	794
Right to reply – applicant has the possibility to read and respond to the reviewer's comment before the final decision	1059	1055
A reviewer is anonymous to the applicant, however, list of reviewers is published after call's financing decisions are made	1655	459

6.5. Have you had experience with a system that allows applicants to nominate possible reviewers?

No	Yes
593	754

6.6. If yes, did you nominate reviewers?

No	Yes

87	667

6.7. Have you had experience with a system that allows applicants to exclude reviewers?

No	Yes
680	636

6.8. If yes, did you give names of potential reviewers to be excluded?

No	Yes
274	362

7.1. What is your opinion about the future of the peer review system?

	Yes
Current system is sufficient	567
There is a need for improvements	848