Authors: Stefan Hornbostel, Meike Siekermann

Peer Review: Healthy to the Core or Chronically Ill?

1. Criticism of the Peer Review System

A competent, neutral and critical peer review system is a pre-condition for meaningful and future-oriented decisions in any peer-review-based process. Since the 1960s, doubts have been voiced periodically regarding the mode and efficiency of review procedures. Some of the criticisms voiced on the peer review system are: The predictive validity is criticised, because some analyses show that reviewer verdicts are not related to subsequent citation rates. Furthermore, some believe that the review system is unreliable, since reviewers sometimes do not agree with each other. In addition, some critics believe that the review method is not valid, because reviewer verdicts are not related to subsequent citation rates. Some studies show that reviewing is biased, since the reviewers' verdicts are influenced by the prestige of the author or institution (cf. Peters and Ceci 1982). Besides, it is also claimed that peer review makes cronyism possible, as the established scientists support each other (Over 1996, Wenneras & Wold 1999). Additionally, some critics accuse the system of discriminating against some groups, such as women and ethnic minorities (cf. Wenneras and Wold 1997). And peer review is judged to be too slow and too expensive (cf. Altman 1996). At least some scientists criticise the fact that peer review could stifle innovation and perpetuate the status quo (cf. Horrobin 1990, Mahoney 1977).

In the context of this criticism, an entire research field has been established. Initially, a study on the American National Science Foundation (NSF) was carried out in 1978. This study, the authors concluded, showed that 50% of the success of an application is determined by coincidental factors connected to the choice of experts and 50% by objective factors: "(…) getting a research grant depends to a significant extent on chance. The degree of disagreement within the population of eligible reviewers is such that whether or not a proposal is funded depends in a large proportion of cases upon which reviewers happen to be selected" (Cole et al. 1981: 881; cf. Cole et al. 1978).

Since then many studies have tried to prove or disprove this criticism. Most peer review studies have concentrated on manuscript reviews. In summary, these studies do not give a clear picture. For example, only a few studies exist that evaluated the predictive validity of decisions together with the criteria of success, and the results are mutually contradictory (e.g. Chapman and McCauley 1993; Armstrong et al. 1997). The research results on the effect of various biases are heterogeneous, too: some studies that analyse gender bias in review processes point out that women scientists are at a disadvantage (e.g. Ferber, Teimann 1980; Wenneras, Wold 1997). However, a similar number of studies merely report moderate or no gender effects (e.g. Gilbert et al. 1994; Jayasinghe, Marsch & Bond 2001). Bornmann et.al. (2006) made a meta-analysis of 21 studies dealing with gender differences in grant award procedures. Even though the estimates varied substantially from study to study, their estimation shows that all in all among grant applicants men have statistically significant greater odds of receiving grants than women by about 7%.

The contrasting research results are also found in the criticism of "old boys networks". Some studies that investigated a positive effect (e.g. Over 1996, Wenneras & Wold 1999), while others did not find any such effect (e.g. Cole 1992). Altogether, some ambiguity exists as to "how 'cronyism' is an artefact of the better reviewers" – the more prestigious the applicant's department is perceived to be the reviewer, the better the rating – but adjusting for the individual track record removed this association" (Wessley & Wood 1999: 17)

Bornstein even reasoned that "peer review fails miserably with respect to every technical criterion for establishing the reliability and validity of an assessment instrument" (Bornstein 1991: 139). This statement does not hold out against empirical analyses. Various studies agree that kappa coefficients between 0.20 and 0.40 correspond to a relatively low level of reviewer agreement (cf. Cicchetti 1991, 1997; Weller 2001: Ch. 6). But there is disagreement on the question of whether higher values would be desirable. Some argue that reviewer disagreement is not a negative factor, and "many see it as a positive method of evaluating a manuscript from a number of different perspectives." (Hans-Dieter Daniel 2005: 145).

2. DFG Reform of the Peer Review System

In autumn 2004, the German Research Foundation (DFG – Deutsche Forschungsgemeinschaft) reformed its peer review system by implementing a new element into the system: the Review Board (Fachkollegium). With the establishment and constitution of its Review Boards, the DFG replaced its previous peer review bodies (Review Committees). Before 2004, the scientific review process was organised into "Review Committees", which were subdivided according to subject. They were elected for each subject on the basis of nominations put forward by the pertinent research societies in a secret ballot of all scientists and academics working in the respective fields at universities and research institutes. Two expert reviewers (Fachgutachter) judged each proposal independently. After considering the two reviewers, the chair of the review committee was responsible for drafting a recommendation, which formed the basis for the decision by the Joint Committee (Hauptausschuss), which is responsible for deciding on research funding and on questions of funding policy.

The increasing number of proposals and the increasing degree of specialisation in scientific disciplines led to numbers of a new expert group, called "special reviewers" (Sondergutachter), increasing. They were originally thought to be more-or-less an exception to the rule. This unelected group supported the "Review Committees" in reaching a funding decision.

While the peer review process was adopted practically unchanged from the predecessor to the current organisation in the early 1950s, the average annual number of proposals was below two thousand. Today, the number of proposals has risen approximately tenfold. Subsequently during the period from 1999 to 2001 almost 9,000 of these "special reviewers" supported the almost 1,000 elected expert reviewers (DFG 2003: 75). This was one reason, among others, for reforming the DFG Peer Review System. Furthermore the International Commission on the Evaluation of the DFG and the Max-Planck-Gesellschaft criticised the structure of the DFG's review committees in 1999 for being too fragmented and outdated (cf. Internationale Kommission zur Systemevaluation 1999).

The main objective of the reform was to increase the influence of the elected review board members, who ensure the scientific quality of the overall review process. This new system leads to greater internal transparency when introducing a new group into the DFG review system in which elected researchers monitor or review the process. The new system was also redesigned in such a way as to respond more quickly to changes in the research system, particularly with regard to interdisciplinary needs and developing new subjects and research areas.

One of the main tasks of the review board members (Fachkollegiaten), who serve in an honorary capacity, is to ensure the quality of the DFG's review process. Now, referees selected by DFG Head Office undertake the task of reviewing the proposal. The board members, elected by the scientific community, are responsible for judging the quality and fairness of the reviews and the reviewer selection. The aim is to achieve a clear distinction between the review, as such, and the overall evaluation of the review process (quality assurance).

3. Survey of DFG Review Board Members

3.1 Methodology and Database

The iFQ conducted a survey of various DFG review board members. The aim of this study was, on the one hand, to obtain information on experience with the reformed DFG review system and, on the other, to identify the strengths and weaknesses of peer review from an expert perspective.

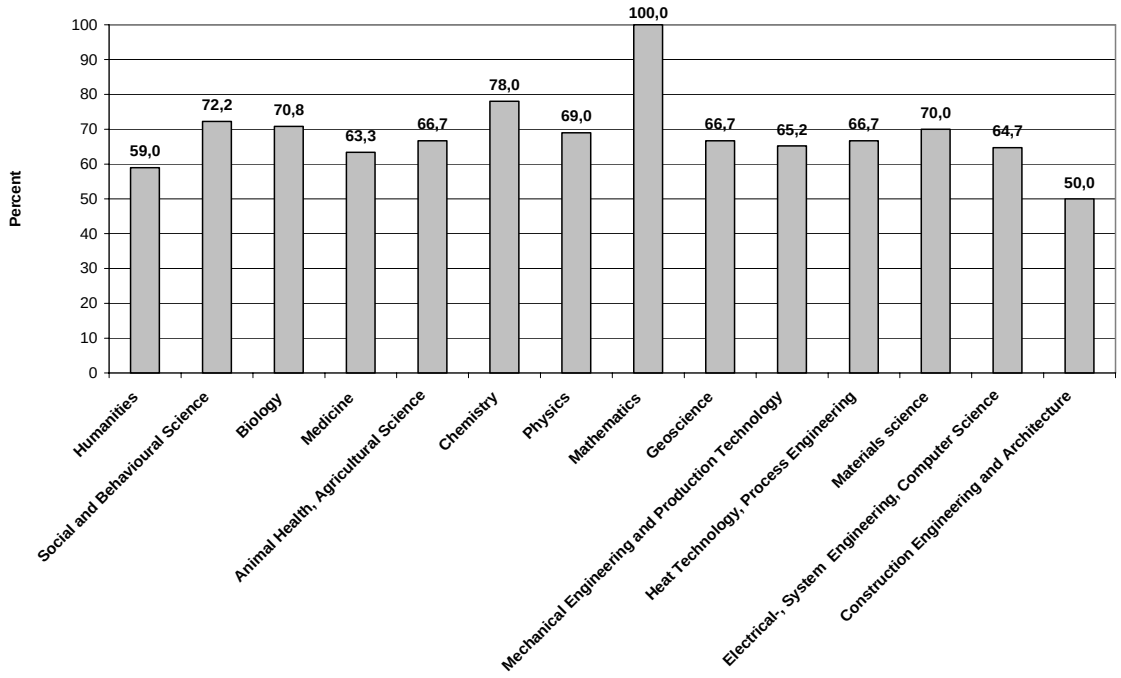
The review board is made up of a total of 577 elected scientists and academics, which in turn work on 48 review boards, which are subdivided into a total of 201 subject areas. This incorporates a total of 14 research areas and finally four scientific disciplines (Humanities and Social Sciences, Life Sciences, Natural Science, Engineering Science).

The survey was carried out in autumn 2006. The link to the online questionnaire was sent to all members of the review boards. The questionnaire, which consisted of 49 mainly multiple choice questions, included the following category groups:

- Experience with the new review board system: have the reform goals been achieved?
- Individual operation of the various review boards: how do they work?
- Reviews and Reviewers: criteria for reviewer selection and the quality of reviews
- Transparency: how important is it?
- Final reports and reviews of final reports: how to handle them in the future?

Furthermore, the questions contained in the Allensbacher Hochschullehrerbefragung (general survey of German professors) of 1976 and 1983 on expert opinions were repeated.

Figure 1: Response rate



457 of the 577 interviewees responded, yielding a rate of return of 79.2%. Figure 1 only represents 66.9% of the interviews, because 71 asked scientists did not mark their subject. The following presents some selected results of this survey.

3.2 Selected Results

3.2.1 Importance of criteria for the selection of reviewers

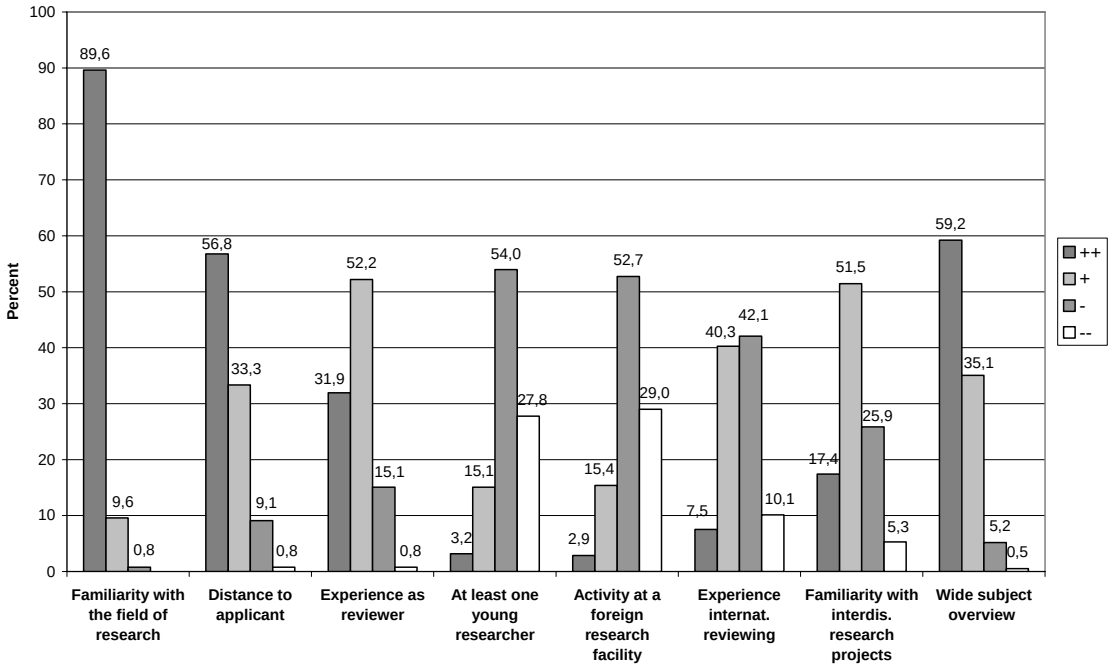
The selection of a reviewer may influence the outcome of a recommendation by the review board. Reviewer selection is therefore one of the key factors and must be done with care. In the DFG system, reviewers are chosen by the DFG programme directors. The review board ensures that appropriate reviewers are selected. One of the questions in the survey asked members of the review boards: How important are various criteria for the selection of reviewers?

It was not astonishing to find that most of the surveyed scientists (89.6%) mentioned that familiarity with the field of research for the proposal is very important for the selection of reviewers. More than half of the interviewees think that a wide subject overview is very important as well (59.2%).
Besides these criteria, around 57% of the members of the review boards believe that distance between the applicant and the reviewer is important.
They do not agree (81.8%) with the statement that at least one young researcher should take part as a reviewer. They are also of the opinion that experience as a reviewer is important. (Cf. Figure 2)
These results are similar to the results produced by Gordon when he interviewed editors of journals. Editors believe that the younger reviewers were too concerned with details

and failed to see the "purpose and significance" (Gordon 1980: 269) of a study. Weller concludes from these results that the age or experience of reviewers might play a role in reviewer performance (cf. Weller 2001: 154).
In contrast to Gordon's results, Evans and colleagues analysed reviews for the Journal General Internal Medicine. Among other things they compared the quality of the reviews with a number of reviewer characteristics. They found that younger reviewers and reviewers from prestigious institutions were more likely to produce better reviews (cf. Evans et al. 1993). The same findings reported Callaham and Tercier (2007) based on an analysis of the "Annals of Emergency Medicine". They reasoned that experienced reviewers perform lower-quality reviews than do younger ones.

Figure 2 illustrates how activities at a foreign research facility plus experience in international peer review processes do not seem to be important for members of the DFG review boards as a selection criteria.

Figure 2: Importance of criteria for the selection of reviewers



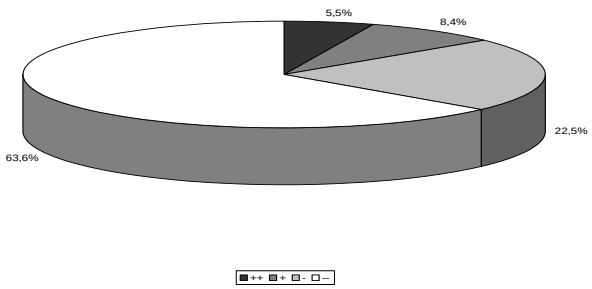
Legend: "Important" corresponds to the category "++" and "unimportant" to "- -"; "+" and "-" are nuances in between

Some scientists used the opportunity to comment in writing on this question. Mostly, they summarised what they believe to be the most important criteria for reviewer selection. The answer given by a scientist from the research area "Electrical and Systems Engineering and Computer Science" serves as a representative example:
"Other very important criteria: 1) Expertise!!! 2) Reviewer must be known for fairness 3) Reviewer must be reliable and work in time."

3.2.2 Transparency

The question of how much transparency is meaningful in the peer review process is often discussed. In practice, experience has been gained with varying degrees of transparency: from total opening as in Denmark (name of reviewer known to the applicant, who can respond to the peer review) to procedures with practically no transparency. And a range of hybrid or mixed procedures exist in between these two extremes.
The anonymity of reviewers ensured that they maintained their critical faculty and their independence towards colleagues. The risk of anonymity is that the process of peer review cannot be controlled or monitored (cf. Hornbostel / Simon 2006).
The question of transparency is closely connected with the degree of anonymity. Various forms of reviewer and author knowledge of each other exist, such as open review (both author and reviewer know each other), anonymous (the reviewer knows the author, but the author does not know the reviewer), or double blind review (neither author nor reviewer know each other).
Most of the members of the review boards (63.6%) do not think that it would make sense for proposals to be anonymous, as illustrated in Figure 3. Only 13.9% hold that double blind procedures are desirable or very desirable.
No big differences exist between the various disciplines when answering this question. With the exception of the "Humanities" (28.2%) and "Electrical and Systems Engineering and Computer Science" (23.8%), which believe that a double blind review would make sense.

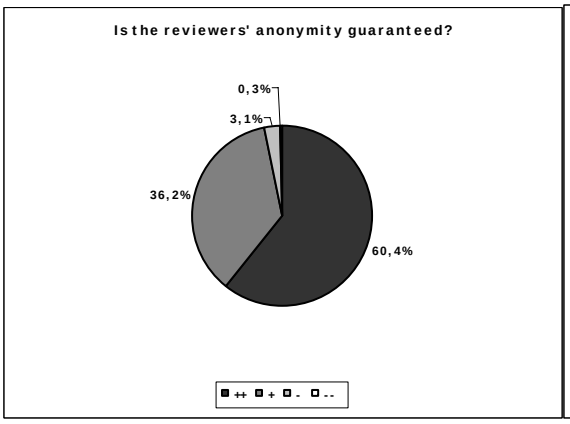
Figure 3: Do you think it would make sense for proposals to be anonymous (double blind review)?



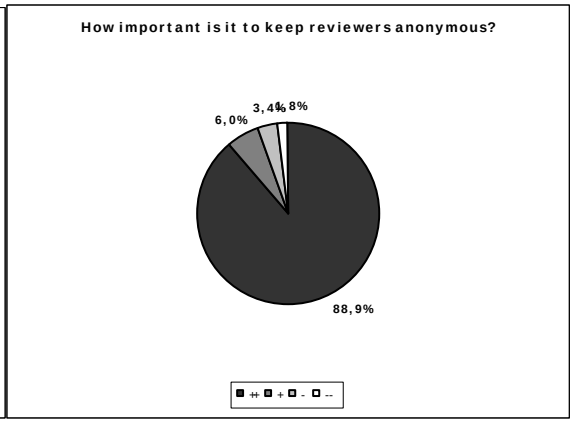
Legend: "Reasonable" corresponds to the category "++" and "Not reasonable" to "- -"; "+" and "-" are nuances in between

When asked about how important it is to keep reviewers anonymous, 88.9% answered "important". And a group of 60.4% thought that the reviewer's anonymity is effectively guaranteed. (Cf. Figure 4)

Figure 4: Questions of anonymity



Legend: "Yes" corresponds to the category "++" and "No" to "- -"; "+" and "-" are nuances in between.



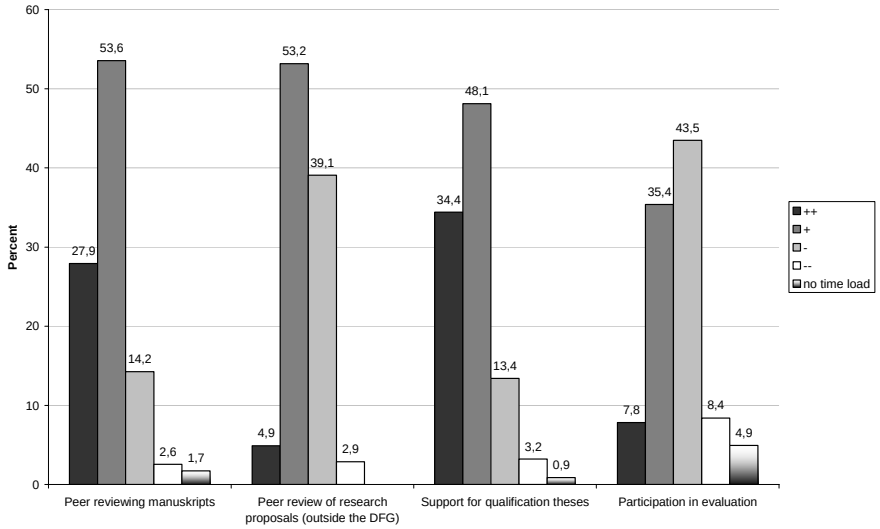
Legend: "Important" corresponds to the category "++" and "unimportant" to "- -"; "+" and "-" are nuances in between.

3.2.3 Time requirement for peer reviewing

Peer reviewing is time-consuming. Peter Roepstoff from the University of Southern Denmark (Odense, Denmark) said: "Peer Reviewing is an integral part of quality assessment in science. For me, however, it has reached a level where the time required for reviewing is beyond my capacity. On average I get 2 to 3 requests to review manuscripts per week and a similar number to review research proposals and reports from research project, plus 1 to 2 requests per month for site visits. This is not manageable, even if I used all my time for this, totally ignoring my own research, my students and my family." (Hans Ulrik Riisgård et al. 2004: 302)

It is to be expected that members of the review boards are in demand. When asked if they regularly do review work outside the DFG, 92.4% confirmed that they do. Reviewing students or PhD theses (qualification work) is very time-consuming. 82.5% of the interviewees specified that the time requirement for such work is high or very high. And the time needed for reviewing journal manuscripts in nearly as high (81.5%), followed by reviewing research proposals that do not come from the DFG (58.1%). Less then half of the interviewees (43.2%) stated that their time requirement for participating in evaluations is high or very high. (Cf. Figure 5)

Figure 5: How high is your time requirement for peer reviewing outside the DFG?

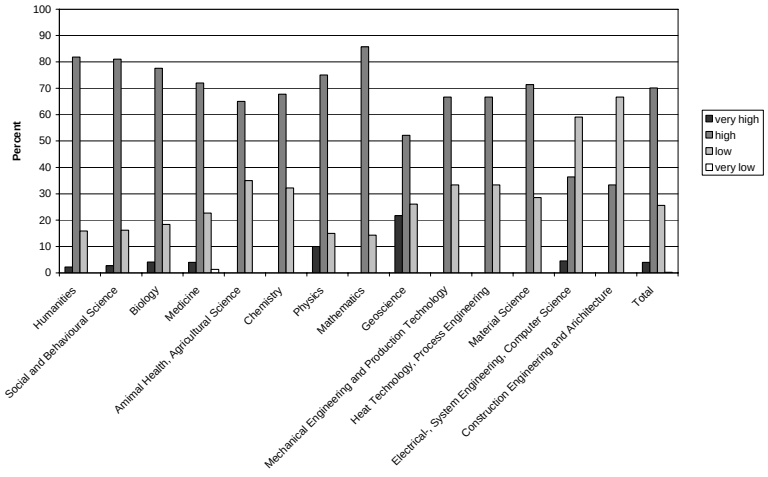


Legend: "High" corresponds to the category "++" and "Low" to "- -"; "+" and "-" are nuances in between.

Besides these forms of peer review, we asked the members of the review boards to estimate the time they spent working as a member of their review board. In total, only 4% of the scientists said that it was very time-consuming (very high), with most of them (70.1%) responding that the time load is high. 25.6% find it to be low and practically no one answered "very low".

Figure 6 illustrates that the research areas "Electrical and Systems Engineering and Computer Science" and "Construction Engineering and Architecture" feel that their time load as members of the review boards is not very high, in contrast to the other areas.

Figure 6: How high is the time requirement as a member of a review board?



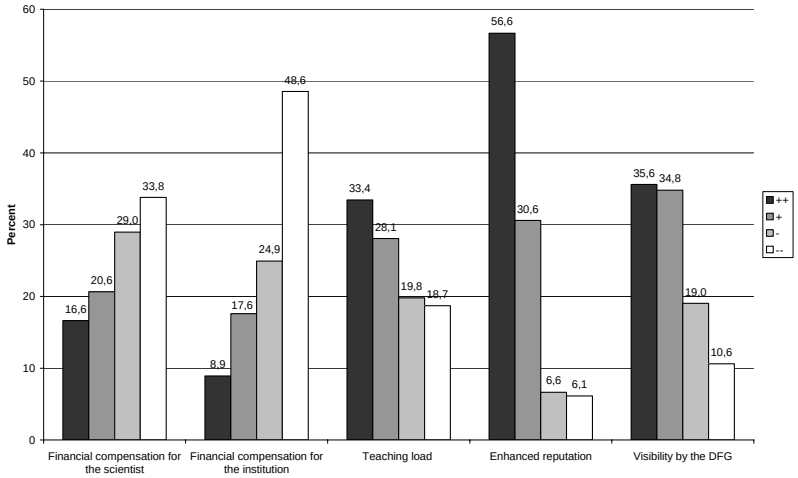
These results show that peer reviewing is in fact a time-consuming job. The review board members serve in an honorary capacity. Therefore, the question of whether the work should be remunerated is often asked. Hans Ulrik Riisgård invited a number of well-established scientists to comment on peer reviewing journal manuscripts in contrast to

research proposals. He came to the conclusion that "(m)ost scientists think that a financial compensation by funding agencies would be fair" (*Hans Ulrik Riisgård et al.* 2004: 309). We asked whether the members of the review boards think that a consideration in return to their work should be made. In contrast to Riisgård's results, they do not think that financial compensation by the DFG would be a good idea. In general 62.8% of the "Fachkollegiaten" disagree with financial compensation, which seems to demonstrate a strong identification with the DFG as a central self-governing organisation of science and research in Germany (cf. Figure 7). But these estimates are not evenly distributed by percentages across all research areas. For example most of the interviewees from "Heat Technology, Process Engineering" (83.3%), "Mechanical Engineering and Production Technology" (80%), "Geoscience" (73.9%) and "Materials Science" (71.4%) did not agree or only partly agreed with the statement "Members of the review boards should be remunerated by the DFG". By contrast, more than half of the interviewed scientists from "Biology" (58.3%) agreed or partly agreed with this statement. The subject areas "Physics" (50%), "Medicine" (43.7%) and "Humanities" (43.4%) also had a comparatively high agreement rate.

The option of the institution at which the scientist works receiving financial compensation was mostly rejected. No major subject-specific differences were found in the answers to this question.

Figure 7 shows that most of the scientists in all subjects agree with the statement: "The role as a member of the review board should be made visible in the reputation of the institution at which I work (e.g. in rankings)." This result demonstrates that reviewers want to be paid in their own "currency", which is "reputation".
The statement: "Members of the review boards should be compensated through teaching loads" also found approval.

Figure 7: Should a consideration be made in return for the work performed as a member of the review board?



Legend: "Yes" corresponds to the category "++" and "no" to "- -"; "+" and "-" are nuances in between

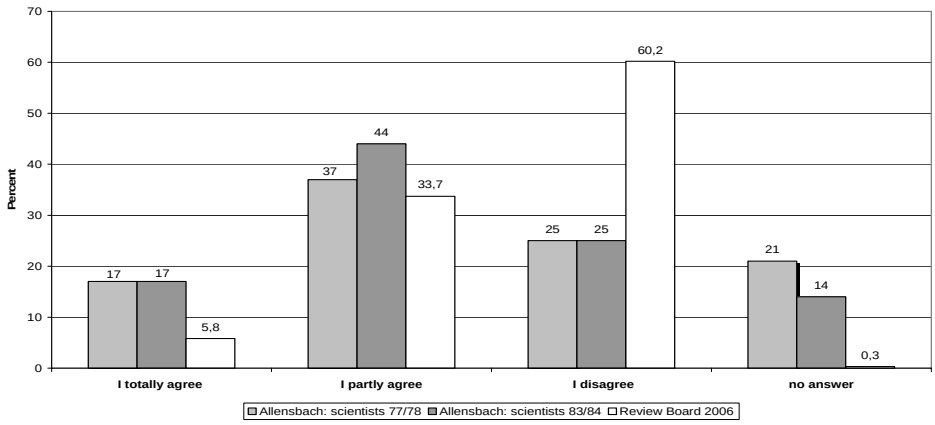
3.2.4 Allensbacher Hochschullehrerbefragung

30 years ago, the public opinion polling institute "Institut für Demoskopie Allensbach" conducted a survey of German professors. Some of these questions dealt with the peer review system of the DFG. The survey was repeated seven years later.

The questionnaire sent to the members of the review boards asked some of these questions again in order to compare differences and identify developments over time. The interviewees could agree or disagree with various statements on a scale of one to four. The results of two of these questions are illustrated below (cf. Figures 8 and 9).

Comparison of the data is not that easy, because overlaps exist between changes in time, differences in the opinions of scientists, in general, and of the members of the review boards, in particular.
Figure 8 illustrates the different answering patterns. While the surveyed professors answered similarly in 1977/78 and 1983/84, some differences were found in the answer pattern in 2006: more than half of the members of the review boards did not agree with the statement "Famous researchers get all their projects – even those that are less excellent – approved". In 1977/1978 and 1983/1984, by contrast, only 25% of the German professors totally disagreed with this statement. In 2006, only 5.8% of the members of the review boards believed that famous researchers are preferred by reviewers. A 33.7% group of review board members remain who partly agree with the statement. These results are not very surprising, because if such a sentence were to attract substantial agreement by the members of the review boards, this would raise the question of whether they are doing their job properly.

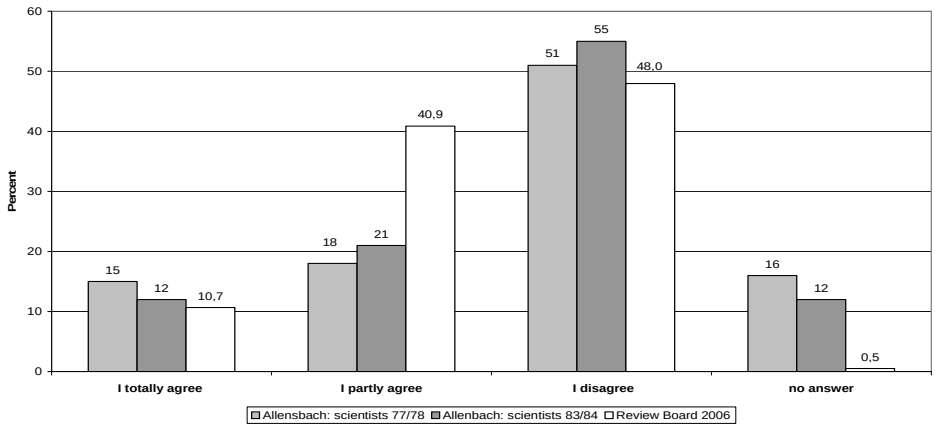
Figure 8: Statement: "Famous researchers get all their projects – even those that are less excellent – approved."



Thirty years ago, discussions were already being held in Germany on the increased use of foreign reviewers.
48% of the interviewed members of the review boards did not agree with the statement "In Germany, there are only a few neutral referees in my research areas. It is necessary to

consult more foreign referees." In 1977/1978 and 1983/1984, the number of scientists who disagreed was only a slightly higher (1977/1978: 51%; 1983/1984: 55%). The number of interviewees in complete agreement with this statement is still low and even lower than it was thirty years ago (1977/1978: 15%; 1983/1984: 12%; 2006: 10.7%). The second big group of interviewed review board members partly agreed with this statement (40.9%). (Cf. Figure 9)

Figure 9: Statement: "In Germany, there are only a few neutral referees in my research areas. It is necessary to consult more foreign referees."

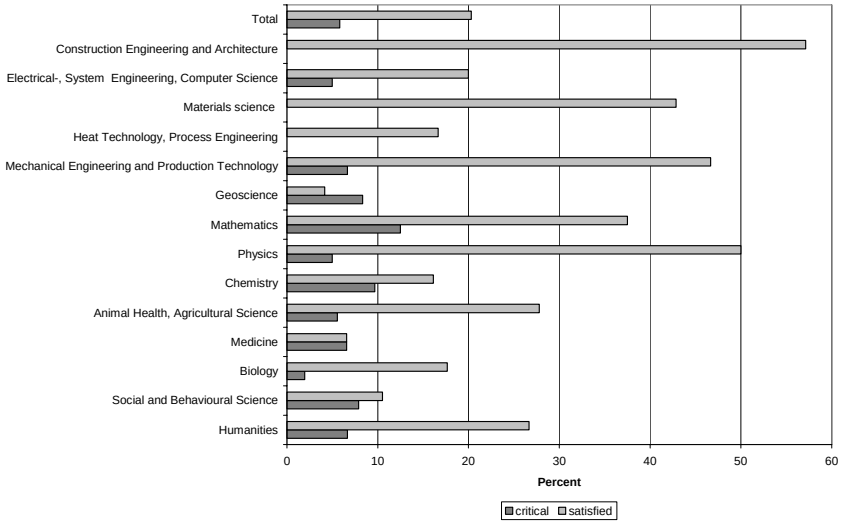


3.2.5. Quality of Reviews

In the written evaluation procedure, the reviews and the proposal are forwarded to the review board. The reviews are one of the basic elements in helping the review board members reach a decision on their recommendation. Ensuring the quality of the reviews is one task of the review boards.

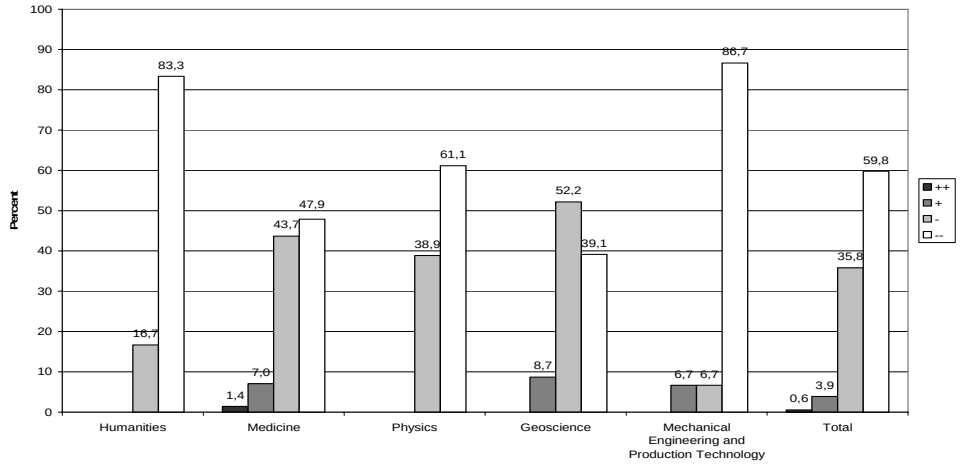
We asked the members of the review boards how high they estimate the percentage of reviews of adequate quality to be. In general, 67.5% of the interviewees answered that 60%-80% of the reviews are of adequate quality. But there are differences between the subject fields. Figure 10 illustrates two groups of review board members from the different research areas: one is a group of critical verdicts on the quality of the reviews, the other a group of satisfied verdicts. Those who said that a percentage of 0% to 50% of the reviews are of adequate quality we defined the "critical" group. Those who said that the percentage of adequate reviews lies between 81% and 100% are the "satisfied" group. The interviewees who answered in between make up the "neutral" group, which is not reflected in Figure 10.
While no one from the "Materials Science" and "Construction Engineering and Architecture" came from the critical group, for example, 12.5% for "Mathematics" and 9.5% for "Chemistry" did belong to this group. In Medicine, the biggest group belongs to the "neutral" section, while "critical" and "satisfied" groups are very similar in size. "Geoscience" is the only subject area where the "critical" group is bigger than the "satisfied" group, but most of the interviewees belong to the "neutral" group here, too.

Figure 10: Quality of reviews: "critical" and "satisfied" verdicts



In contrast to the question on the quality of reviews, in general, the answer to the question on the quality of foreign reviewers depends heavily on the subject. Figure 11 illustrates the typical answering pattern in some subject areas. In total, more than half of the interviewees did not think that the quality of foreign reviewers is better than that of national reviewers. The percentage of interviewed scientists who think that the quality is better is negligible (0.6%). Members of the review boards who belong to the subject area "Mechanical Engineering and Production Technology" completely reject this statement with a value of 86.7%. A high degree of disagreement is also found in "Humanities" and "Physics". Not even one interviewee from these two subject areas supports the statement. The view in the subject area "Geoscience" is slightly different. Nobody here agrees with the statement, but there is smaller group who totally disagree (39.1%) and a bigger group who disagree somewhat (52.2%). "Medicine" also has a small group who think that the quality of foreign reviewers is better or somewhat better (8.4%). (Cf. Figure 11)

Figure 11: Statement: "The quality of foreign reviewers is better than that of national reviewers."



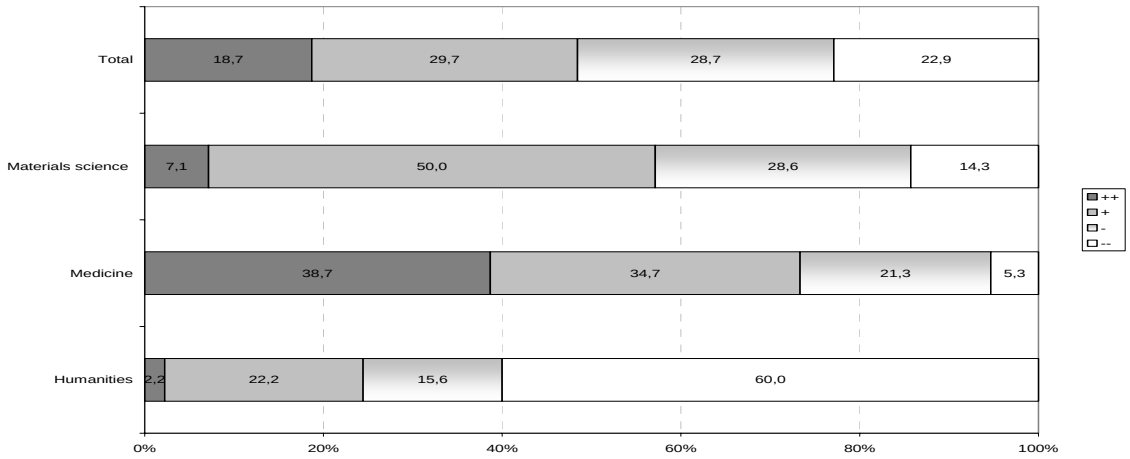
Legend: “I agree” corresponds to the category “++” and “I do not agree” to “- -“; “+” and “-“ are nuances in between

3.2.6 Citation Index

"Informed peer review" is an extended form of peer reviewing: the reviewer's opinion forming is supported by quantitative indicators. Hence, this form of review links qualitative research indicators, like the citation index of publications, with the reviewing of peers. We asked the board members whether using the citation index would be helpful in the review process.

On average, one in two of the interviewed board members stated that they do not or only slightly believe that the citation index of publications would be helpful in the review process.

Figure 12: Would the citation index of publications by the applicants be helpful in reaching a decision?



Legend: "Yes" corresponds to the category "++" and "no" to "- -"; "+" and "-" are nuances in between

Differences can be seen between the disciplines: 73.4% in medicine would find a citation index helpful. In Materials Science, 57.1% of the board members still have a positive attitude to using the citation index. Contrary to these subject areas, most of the members of the review boards who come from the field of humanities do not think that the citation index would be helpful (cf. Figure 12). This is an unsurprising result, because the use of citation indices is not very common and is disputed in the humanities

4. Conclusion

Altogether, the review processes reveal substantial differences between scientific fields. But across all subjects, consensus exists that forms of open review (reviewer known) are not recommended for funding decisions. Conversely, strong recommendations exist for open access to the findings of funded projects.

The interviewees believe that the time load for members of the review boards is time-consuming. But they would rather be rewarded in their own "currency", that is "reputation".

Review board members obviously regard the peer review system in Germany as being internationally comparable. Therefore, they do not see any special advantages in recruit-

ing foreign reviewers. Additionally, they are critical in their judgement of the quality of reviews.
All in all, the new review system seems to be quite reasonable, because the elected board members are critical in judging the quality of reviews. They are sceptical about the quality and, at the same time, show solidarity with the referees, whose anonymity they want to preserve.
Altogether and despite the numerous problems associated with peer review, this system is nevertheless irreplaceable: no one but peers will be able to judge the originality or the degree of innovation of a research proposal. Drummond Rennie, deputy editor of *JAMA*, put it this way in an article summarising the pluses and minuses of the system. "Peer review is like democracy," he wrote, "which is, to use Churchill's phrase, the worst form of government except all those other forms that have been tried from time to time." (Rennie 1993) Establishing a review board seems to offer a good compromise between preserving the anonymity of the reviewers and meeting the interests of the public community.

Literature:

Armstrong, P. W., et al. (1997): Evaluation of the Heart and Stroke Foundation of Canada Research Scholarship Program: Research productivity and impact. Canadian Journal of Cardiology 15(5): 507–516.

Altman, L.K. (1996): The Ingelfinger rule, embargoes, and journal peer review – part 1: Lancet, 347:1382–6; part 2: 347: 1459–63.

Bornmann, L. / Mutz, R. / Daniel, H.-D. (2006): Gender differences in grant peer review. A meta-analysis. http://arxiv.org/abs/math.ST/0701537

Bornstein, R.F. (1991): The predictive validity of peer review – a neglected issue. Behavioral and Brain Sciences, 14 (1), 138-139.

Callaham, M.L. / Tercier, J. (2007): The relationship of previous training and experience of Journal Peer Reviewers to subsequent Review Quality. PLoS Medcine, Vol.4, Issue 1: 32 – 40.

Chapman, G.B. and McCauley, C. (1993): Early career achievements of National Science Foundation (NSF) graduate applicants: Looking for Pygmalion and Galatea effects on NSF winners. Journal of Applied Psychology, 78: 815-820.

Cicchetti, D.V. (1991): The reliability of peer review for manuscript and grant submissions: a cross-disciplinary investigation. Behavioral and Brain Sciences, 14: 119-135.

Cicchetti, D.V. (1997): Referees, editors, and publication practices: improving the reliability and usefulness of the peer review system. Science and Engineering Ethics, 3: 51-62.

Cole, S. / Rubin, L. / Cole, J.R. (1978): Peer Review in the National Science Foundation. Phase I of a Study. Washington: National Academy of Sciences.

Cole, S/ Cole, J.R. / Simon, G. (1981): Chance and consensus in peer review. In: Science. 214: 881-886.

Cole, S. (1992): Making Science. Between Nature and Society. Cambridge: Harvard University Press 1992.

Daniel, H.-D. (2005): Publications as a measure of scientific advancement and of scientists' productivity. Learned Publishing 18: 143-148.

Deutsche Forschungsgemeinschaft (DFG) (1999): Stellungnahme zum Bericht der internationalen Kommission zur Systemevaluation der Deutschen Forschungsgemeinschaft und der Max-Planck-Gesellschaft, Oktober 1999.

Deutsche Forschungsgemeinschaft (DFG) (2003): Funding Ranking 2003. Institutions – Regions – Networks, Bonn 2003.

Evans, A.T., McNutt, R.A., Fletcher, S.W. and Fletcher, R.W. (1993, August): The characteristics of peer reviewers who produce good-quality reviews. Journal of General Internal Medicine, 8 (8): 422-428.

Ferber, M.A. and Teiman, M. (1980): Are women economists at a disadvantage in publishing journal articles? Eastern Economic Journal, VI (3-4): 189-193.

Gilbert, J.R., Williams, E.S. and Lundberg, G.D. (1994): Is there gender bias in JAMA's peer-review process. JAMA Journal of the American Medical Association, 272(2): 139–42.

Gordan, M.D. (1980): The role of referees in scientific communication. In: J. Hartley (ed.), The psychology of written communication, selected readings (pp. 263-275). New York: Nichols Publishing Co.

Hornbostel, S. / Simon, D. (Hg.) (2006): Wie vie (In-)Transparenz ist notwendig? Peer Review revisited. iFQ-Working paper No. 1
http://www.forschungsinfo.de/Publikationen/Download/working_paper_1_2006.pdf

Horrobin, David F (1990): 'The philosophical basis of peer review and the suppression of innovation', Journal of the American Medical Association, 263(10): 1438-1441.

Institut für Demoskopie (1984): Zur Lage der Forschung an deutschen Universitäten, 1977-1984. Tabellenband. Ms., Allensbach.

Internationale Kommission zur Systemevaluation der Deutschen Forschungsgemeinschaft und der Max-Planck-Gesellschaft (1999), Hannover.

Jajasinghe, U.W / Marsh H.W. / Bond, N (2001): Peer Review in the Funding of Research in Higher Education: The Australian Experience. Educational Evaluation and Policy Analysis, 23 (4): 343-364.

Mahoney, M.J. (1997): Publication prejudices: an experimental study of confirmatory bias in the peer review system. Cogn Ther Res 1: 161–75.

Over, R. (1996): Perceptions of the Australian Research Council Large Grants Scheme: Differences between Successful and Unsuccessful Applicants. Australien Educational Researcher, 23 (2): 17-36.

Peters, D.P., Ceci, S.J. (1982): Peer-review practices of psychological journals: the fate of published articles, submitted again. Behavioral and Brain Sciences, 5: 187-195.

Riisgård, H.U. (2004): Peer review: journal articles versus research proposals. Marine Ecology Progress Series 277: 301-309.

Rennie, D., (1993): "More Peering into Editorial Peer Review," Journal of the American Medical Association, 270 (Dec. 15): 2856-2858.

Weller, A. C. (2001): Editorial Peer Review. Its Strengths and Weaknesses. New Jersey: Information Today Inc.

Wennerars, C. / Wold, A. (1997): Nepotism and sexism in peer-review. In: Nature, No. 387: 341-343.

Wenneras, C. and Wold, Agnes (1999): "Bias in peer review of research proposals" in Godlee and Jefferson (Eds.) Peer Review in Health Sciences. London: BMJ Publishing Group: 79-89.

Wessley, S. / Wood, F. (1999): Peer Review of Grant Applications: A systematic Review. In: F. Godlee; T. Jefferson (eds.), Peer Review in Health Sciences, London: 14-31.