

Editorial

Reviewing in science requires quality criteria and professional reviewers

Karin Jurkat-Rott, Frank Lehmann-Horn

Abteilung für Angewandte Physiologie, Universität Ulm, Ulm, Germany

The use of peers to assess the work of scientists goes back to the 17th century and finally led to what we call peer review (Kronick, 1990) or review by competitors (Roy and Ashburn, 2001). Usually peers are external experts not paid for their assessment. Recent studies of the effectiveness and quality of our current peer review system are not reassuring (Godlee and Jefferson, 1999; Jefferson et al., 2002a, b). The continued use rather results from the lack of a serious alternative than from its actual value (Young, 2003). The purpose of our editorial is to address the problems brought up by peer review and their requirements for higher quality. We propose professional reviewing as an alternative to be tested.

Problem 1: reviews call on a scientist's time

In addition to the time taken for research, teaching, and administration, scientists spend approximately 20% of their total working time doing peer reviews: promotional reviews requested by university deans, reviews of applications for project grants or for scientific prizes required by governmental and private research foundations, and, above all, manuscript reviews for editors of scientific journals. This intensive refereeing activity means that scientists must do their usual work in the residual 80% of their time – leading to reduced quality of research and teaching.

Problem 2: peers have conflicts of interests

By doing peer reviews, scientists have the power to determine the success or doom of their competitors, e.g. who will be promoted, have their scientific manuscripts accepted for publication, be granted research money, or obtain a prize.

Corresponding authors: Dr. Karin Jurkat-Rott, Prof. Dr. Frank Lehmann-Horn, Abteilung für Angewandte Physiologie, Universität Ulm, Albert-Einstein-Allee 11, D-89069 Ulm, Germany, E-mail: frank.lehmann-horn@medizin.uni-ulm.de, karin.jurkat-rott@medizin.uni-ulm.de, Fax: +49 731 50023260.

They also gain access to confidential information on newest research results, novel techniques, or innovative ideas. It is not surprising that misuse of these opportunities is frequent. The degree of misconduct ranges from deliberately postponing the review over mild forms of plagiarism to subreption of patent rights and commercial advantages (Dalton, 2001).

Problem 3: many peer reviews are of unreliable quality

There are no standardized quality criteria that peer reviews must fulfill (Kennedy, 2004). Scientists do the reviews with more or less time and devotion, but occasionally with quite a bit of emotion. The reviews may therefore be unsystematic, superficial, uncritical, or biased (Wechsler and Fried, 2003). Even top-ranking journals are not excluded from low quality reviews as several examples of overlooking of the same fabricated or falsified data in several articles have shown (Sellke, 2003; Connerade, 2004). Publication of obviously wrong results or unjustified conclusions frequently occurs. In many journals, post-publication peer review is offered to the discussion of questionable material previously published in the same journal (e.g. (Sood et al., 1996) criticized by Wang et al. (1997); (Ahmad et al., 1998) criticized by Hillmer et al. (2001)) whereas in other cases criticisms may be accepted by other journals (e.g. (Wang and Goldstein, 1995) criticized by Barhanin et al. (1996); (Abbott et al., 2001) criticized by Jurkat-Rott and Lehmann-Horn (2004)). But it is not the extreme cases that are the major problems, rather it is the general quality of reviews. A study on review quality examined the rate of detection of deliberate inconsistencies incorporated into the manuscripts: the majority of the reviewers detected less than 50% of the errors (Godlee et al., 1998). Another study compared the rankings of peers and found a very low inter-rater reliability (Howard and Wilkinson, 1998).

Previous problem-solving strategies have failed to improve review quality

Various efforts to improve peer reviewing have been made in the past. Studies focused on the effects of concealing author identity, a not very successful endeavor (Justice et al., 1998; Van Rooyen et al., 1999) because the reviewers correctly guessed author identity in up to 46% of cases and consequently gave better scores on the work of authors with longer publication records (Fisher and Friedman, 1994; Fisher et al., 1994). Studies of open refereeing, i.e. revealing the identity of the reviewers to the authors, did not show improved quality, either (Godlee, 2002). The use of minimal requirement schemes for reviews or of written feedback on review quality by the editor were both equally ineffective measures (Callaham et al., 2002). Also, attendance of peer reviewers at a highly structured and interactive workshop did not improve the quality of subsequent reviews. Efforts to aggressively recruit average reviewers to a second workshop showed a similar lack of effect on ratings (Callaham and Schriger, 2002). Thus, intrinsic motivation is a prerequisite for excellent reviewing. Until now, only a single review parameter was clearly influenced by strategies: preceding inquiry of potential reviewers for their consent to review the manuscript shortened the time taken to produce the review but, on the other hand, reduced the number of reviewers willing to do the job (Godlee, 2002).

Reviews in other disciplines

In sample areas, reviews by qualified personnel are required for decision making. This is true for judging the worth of an art object or the damage to it, determining the degree of physical disability of the handicapped, legally empowered certifying validity of documents, or making a prognosis with respect to economical or ecological developments prior to investments. These tasks are performed by professionals, i.e. individuals who make their money by selling their expertise in a field. Such professionals cannot afford to be biased otherwise none would buy their services. For the same reasons, they take special care to produce reliable high-quality reviews in due time and to protect the interests of their customers. Therefore, the introduction of a professional refereeing system into sciences is likely to have the same advantages as in other disciplines and lead to increased neutrality, improved quality, avoidance of attacks on the personal integrity of the authors, and more timeliness of the reviews, and, in the end, relief of scientists' work load for the benefit of research and teaching.

Requirements for professional reviewers

As in other disciplines, professional science reviewers would need to be highly qualified. Obvious prerequisites are a university degree with PhD and personal experience in writing of manuscripts and grant applications in the particular field of science to be reviewed (e.g. cell biology). Also, a university certificate on *Good Referee Practice* from a postgraduate vocational training unit would be essential. Regular attendance

of scientific meetings both topical as well as specialized on refereeing such as the International Congress on Peer Review and Biomedical Publication are recommended. Since professionals are intrinsically interested in constantly improving their quality to obtain more commissions, these recommendations will doubtlessly be followed. Not to forget, the constant refereeing activity will ensure that the professional reviewers are indeed up to date and highly experienced in detecting flaws in concept and conclusion of the type of science they are dealing with. Professional reviewers thus defined may not necessarily work as such all their life, it is highly likely that several may spend just a few years after their postdoctoral phase doing this task before moving on to work for research foundations, research lobbies, international research organizations, or publishing houses.

Requirements for manuscripts and applications in science

Due to increasing specialization of the scientists, the expertise even within one research field is extremely limited. In some fields, less than half a dozen research groups around the world are actually working on related topics and are thus considered to be sufficient insiders or, explicitly, experts acceptable for refereeing. General opinion is that no outsider could ever succeed to grasp their work much less judge its quality or make suggestions for its improvement. With this in mind, it becomes clear that the reviewers are not always the problem, but rather the manuscripts and grant applications to be reviewed: they do not fit in the spectrum of scientific work which the general scientific public can comprehend and appreciate – and it is part of the reviewers' job to fulfill the specific quality criteria. Several research funding organizations have responded to these demands and have developed review criteria and effective review procedures to ensure quality and reduce bias and injustice. Examples are not only big federal research foundations like the NIH but even smaller private ones such as Telethon in Italy specialized on funding research into muscular disorders. However, editors of many scientific journals, even top-ranking ones, do not see a necessity for requiring standards for publications or for criteria and procedures of their review process. One important reason could be that such requirements may jeopardize the commitment of their refereeing scientists who are currently doing their job absolutely without gratification.

Are professional reviewers affordable?

This question is actually not the question to be asked. Instead, the real question is: who is currently paying for the reviews that the publishing companies profit from by selling their journals to all university libraries around the world? The answer is that most refereeing scientists are either being paid by federal institutions to do research and teaching or by grant money from research foundations restricted to a specific research project. In both cases, reviewing activities are not included in the scope. As such a situation indeed indicates misuse of public funds, general consensus is that the reviewers are, of course, doing the reviews in their free time for the honor of being appointed and for the

favorable appearance on the scientific curriculum vitae. In reality, most scientists are voluntarily reviewing manuscripts because they see it as part of their scientific duty to maintain the peer review system by their contribution – and therefore consider reviewing as part of their working responsibilities.

Conclusions

Big areas of research and thus the vast majority of publication manuscripts and grant applications are suitable for reviewing by certified professionals. Also, professional refereeing would solve the current conflict of misuse of funds, especially for profit journals which, in contrast to some non-profit journals, do not re-invest the surplus for supporting science.

Authors

HD Dr. Karin Jurkat-Rott is university teacher of Physiology at Ulm University and leader of the “Electric Current & Genetics Group”. Prof. Dr. h. c. Frank Lehmann-Horn is a full professor of Physiology at Ulm University (<http://physiologie.uni-ulm.de>). They have established a module *Good Referee Practice* in the new Master of Science programme on Molecular Medicine at the Medical Faculty of Ulm University. (<http://www.uni-ulm.de/mm/>). The module includes critical reading of and commenting on scientific texts, basics of auditing, and biostatistics; knowledge of good scientific practice, of national and international research structures, of political frameworks, and of committees supplying advice on the requirements on the verification of results and on publication ethics; and attendance of scientific meetings, methodological training courses, and specific congresses on reviewing and publishing.

References

- Abbott, G. W., Butler, M. H., Bendahhou, S., Dalakas, M. C., Ptáček, L. J., Goldstein, S. A. (2001): MiRP2 forms potassium channels in skeletal muscle with Kv3.4 and is associated with periodic paralysis. *Cell* **104**, 217–231.
- Ahmad, W., Irvine, A. D., Lam, H., Buckley, C., Bingham, E. A., Panteleyev, A. A., Ahmad, M., McGrath, J. A., Christiano, A. M. (1998): A missense mutation in the zinc-finger domain of the human hairless gene underlies congenital atrichia in a family of Irish travellers. *Am. J. Hum. Genet.* **63**, 984–991.
- Barhanin, J., Lesage, F., Guillemare, E., Fink, M., Lazdunski, M., Romey, G. (1996): K(V)LQT1 and IsK (minK) proteins associate to form the I(Ks) cardiac potassium current. *Nature* **7**, 78–80.
- Callaham, M. L., Knopp, R. K., Gallagher, E. J. (2002): Effect of written feedback by editors on quality of reviews: two randomized trials. *JAMA* **287**, 2781–2783.
- Callaham, M. L., Schriger, D. L. (2002): Effect of structured workshop training on subsequent performance of journal peer reviewers. *Ann. Emerg. Med.* **40**, 323–328.
- Connerade, J. P. (2004): Scandals stem from the low priority of peer review. *Nature* **427**, 196.
- Dalton, R. (2001): Peers under pressure. *Nature* **413**, 102–104.
- Fisher, M., Friedman, S. B. (1994): The blind reading the blind. *J. Dev. Behav. Pediatr.* **15**, 386–387.
- Fisher, M., Friedman, S. B., Strauss, B. (1994): The effects of blinding on acceptance of research papers by peer review. *JAMA* **272**, 143–146.
- Godlee, F. (2002): Making reviewers visible: openness, accountability, and credit. *JAMA* **287**, 2762–2765.
- Godlee, F., Jefferson, T. (1999): Peer Review in Health Sciences. BMJ Books, London.
- Godlee, F., Gale, C. R., Martyn, C. N. (1998): Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. *JAMA* **280**, 237–240.
- Hillmer, A. M., Kruse, R., Betz, R. C., Schumacher, J., Heyn, U., Propping, P., Nothen, M. M., Cichon, S. (2001): Variant 1859G → A (Arg620Gln) of the “hairless” gene: absence of association with papular atrichia or androgenic alopecia. *Am. J. Hum. Genet.* **69**, 235–237.
- Howard, L., Wilkinson, G. (1998): Peer review and editorial decision-making. *Br. J. Psychiatry* **173**, 110–113.
- Jefferson, T., Wager, E., Davidoff, F. (2002a): Measuring the quality of editorial peer review. *JAMA* **287**, 2786–2790.
- Jefferson, T., Alderson, P., Wager, E., Davidoff, F. (2002b): Effects of editorial peer review: a systematic review. *JAMA* **287**, 2784–2786.
- Jurkat-Rott, K., Lehmann-Horn, F. (2004): Periodic paralysis mutation MiRP2-R83H in controls: interpretations and general recommendation. *Neurology* **62**, 1012–1015.
- Justice, A. C., Cho, M. K., Winker, M. A., Berlin, J. A., Rennie, D. (1998): Does masking author identity improve peer review quality? A randomized controlled trial. *PEER Investigators. JAMA* **280**, 240–242.
- Kennedy, D. (2004): Disclosure and disinterest. *Science* **303**, 15.
- Kronick, D. A. (1990): Peer review in 18th-century scientific journalism. *JAMA* **263**, 1321–1322.
- Roy, R., Ashburn, J. R. (2001): The perils of peer review. *Nature* **414**, 393–394.
- Sellke, F. W. (2003): The peer-review process in medical publishing: a reviewer’s perspective. *J. Thorac. Cardiovasc. Surg.* **126**, 1683–1685.
- Sood, S., Eldadah, Z. A., Krause, W. L., McIntosh, I., Dietz, H. C. (1996): Mutation in fibrillin-1 and the Marfanoid-craniosynostosis (Shprintzen-Goldberg) syndrome. *Nat. Genet.* **12**, 209–211.
- Van Rooyen, S., Godlee, F., Evans, S., Smith, R., Black, N. (1999): Effect of blinding and unmasking on the quality of peer review. *J. Gen. Intern. Med.* **14**, 622–624.
- Wang, K. W., Goldstein, S. A. (1995): Subunit composition of minK potassium channels. *Neuron* **14**, 1303–1309.
- Wang, M., Mathews, K. R., Imaizumi, K., Beiraghi, S., Blumberg, B., Scheuner, M., Graham, J. M., Jr., Godfrey, M. (1997): P1148A in fibrillin-1 is not a mutation anymore. *Nat. Genet.* **15**, 12.
- Wechsler, A. S., Fried, P. W., 2003. Peer review. *J. Thorac. Cardiovasc. Surg.* **126**, 1681–1682.
- Young, S. N. (2003): Peer review of manuscripts: theory and practice. *J. Psychiatry Neurosci.* **28**, 327–330.