

# Peering into review

Rajendra Persaud

In his recent interview with Greg Wilkinson, past-president Jim Birley dryly reminds us that his first paper on life events was turned down by the *British Journal of Psychiatry*, as it was then thought too far-fetched that schizophrenia could be precipitated by a life event (Wilkinson, 1995). Thankfully fashions in psychiatry, as in many fields, have changed since the 1960s, but as science is meant to be concerned with more enduring truths, this throwaway comment raises worrying questions. How much valuable but 'unfashionable' research is today being rejected by our journals?

Psychological and psychiatric factors among researchers, like prejudice, fear and even jealousy, are, it is hoped, removed from the selection of papers for publication by the widely respected system of peer review. However, there is accumulating evidence that there are fundamental weaknesses in the peer review process.

Peters & Ceci (1982) took papers that had already been accepted by behavioural science periodicals, and resubmitted them to the same journals in which they had been published, but after removing the prestigious departments and authors from the title pages. Only one quarter of the published articles were even recognised by the reviewers as having already been published, and 88% of the rest were then rejected as unsuitable for publication by reviewers of the same journal in which they had already been published.

Other research has documented very low reliabilities and agreement between referees (Marsh & Ball, 1989). If single-reviewer reliability is defined as the correlation between two independent reviews of the same manuscript across a large number of manuscripts submitted for publication, then Marsh & Ball (1989) found single-reviewer reliability was a mean of only 0.27 across ten behavioural science journals. Given the vast care and resources invested in the production of papers, it appears an irrational part of the scientific process that the crucial element of

peer review is so neglected (Ingelfinger, 1974). This is particularly intriguing given the recent spate of published reports concerned with scientific fraud (Atkinson, 1994). Yet unreasonable opposition from referees which suppresses a new idea could be seen as unwelcome to science as fraud. While fraudulent claims will not survive replication, a suppressed idea can be lost forever.

## Practical problems, practical solutions

A survey of the US-based journals with the highest scores on the 1989 Scientific Information Citation Frequency Index reveals that these journals leave most reviewers with vague global instructions; for example, to assess the paper for its suitability for publication (Cotton, 1993).

Editors should make explicit that judgements are required of reviewers on methodological and non-methodological elements of a paper. Non-methodological issues include clarity, novelty, clinical or theoretical importance, and whether the conclusions are reasonable.

In the assessment of the method a reviewer should determine carefully and separately groundedness (ignoring important work of others); samples (representative, large enough, inadequate controls); measurements (incorrect, imprecise, neglected or clinically unimportant instruments); experimental design (inadequate follow-up, confounding variables, cause and effect not established); data (inappropriate manipulation, unsuitable statistical tests).

Referees should *advise* editors, not *decide* for them. All referee recommendations or comments should have the reasoning behind them clearly elaborated, and editors should give authors an opportunity to reply to a reviewer's negative comments before reaching an editorial decision, especially when the reviews are mixed. This would minimise the problem of having a quality paper rejected outright because some reviewers are uninformed, overly demanding, or

misunderstand the paper. The reviewers in turn could be invited to respond to the author's reply if desired. Editors would then be able to make a final decision based on a fair balance of opinions (Finke, 1990).

### Alternatives to peer review

So far these arguments are based on the assumption that peer review done well would represent a scientific ideal. However, it might be that the whole concept of peer review is in itself flawed.

One study found that between 68% and 99% of papers published in a sample of highly reputable journals had to be revised subsequent to submission (Bradley, 1982). These high rates of coerced revision emphasise the enormous power of referees to enforce their views.

The knowledge explosion, increasingly complex methods, and the narrowing of specialisation mean that identifying true scientific peers for a particular topic under review is problematic. The usual resort is to consult the eminent, but there are numerous problems with this approach. No less than Darwin and Planck experienced the resistance of established scientists to questioning accepted paradigms (Kuhn, 1970, pp 150–153). In fact intimate association with a research area could adversely effect impartial judgement. The longer the engagement with a particular research question, perhaps the greater the personal involvement.

The history of peer review, at first glance, appears honourable, reflecting the esteem with which it is held by scientists. In the 17th century the Royal Society of London established a board of editors to evaluate reports submitted for publication in its *Proceedings*. However, the true birth of peer review could be claimed by the Jesuit astronomers who dismissed Galileo's revolutionary astronomical claims. Spectacular failures of peer review are rarely considered by its defenders, who claim any weaknesses are acceptable failings given its strengths.

Journals are not just the dry receptacles for work by automatons. All papers are the end symbol of an important enterprise, the actual physical work of a group of individuals. Rejection of a paper condemns a large amount of human activity as having been a waste of time. This contributes to the

competitiveness and stress of academic life. But surely all research work has some good qualities and some bad, and a better use of peer review could be to consider it almost as a behavioural shaping exercise – encouraging good work, identifying and discouraging poor methods.

Using this framework no paper should ever be rejected outright. Instead what is good should be identified and praised, what is bad should be distinguished and suggestions made to modify, so signposting the way to improve the paper. This means transforming peer review from an intimidating hurdle to something more like an encouraging tutorial. There are several possible benefits; paper submission could become a less disheartening prospect to psychiatrists just starting out on research; all reviews would have to be undertaken more carefully if the good as well as the bad aspects of a paper have to be identified, and this form of review would make it easier to detect the negative review being written for less than objective reasons.

Obviously if all papers are considered ultimately publishable, the only difference between poor and good work would be the much longer delays that initially inadequate research would experience before seeing the light of day, due to the extensive revision required or referral to other more suitable journals. Most authors respond well to constructive criticism or advice, especially if it is required to achieve publication (O'Connor, 1978).

Perhaps one reason peer review has not been adequately subject to review itself is the implicit feeling that the voluntary work of the reviewers should not be directed too closely or criticised. Given these problems, some have suggested the privilege of high quality feedback on our work should be worth paying for, in the form of a submission fee charged to an author for sending a paper to a journal (Bornstein, 1990). Obviously this would raise difficulties for those attempting to submit papers on third world research budgets, but special considerations could apply in these situations. Most importantly, it is possible that more thoughtful and reasoned reviews would result from making this a lucrative part of academic life.

These perspectives remind us that publication is not an end in itself, instead publication enables debate, leading to attempts at replication, and then archiving.

An implicit assumption of peer review is that deliberation over a paper only starts with publication, yet it is clear that debate begins well before this stage, and starts with peer review itself.

This is recognised by those journals which practise open peer commentary, in which the peer review is published with the paper. While this allows the peers' judgements and thought to be scrutinised as much as the paper itself, it still avoids the problem of deciding who qualifies as a suitable peer.

Published correspondence is more genuinely peer review because it is likely to generate further correspondence, ensuring the peer review is itself peer reviewed. Correspondence is usually spontaneously elicited from the genuinely interested, rather than just the eminent.

It is perhaps sobering to realise that after publication less than 1% of papers in medical journals are read through to the end (Cotton, 1993). This suggests that even if the current system of peer review is considered good enough, it is delivering papers for publication which do not attract much interest. If journals are serious in their intention to communicate widely progressive, novel and therefore perhaps controversial ideas, it might be that the onus of proof in a disagreement should sometimes be firmly on the shoulders of critical reviewers. Otherwise journals will tend to preserve orthodoxies rather than advance debate, and the vast majority of clinicians may see little need to read journals.

### Comment

Peer review is only as good as its peers, so maintaining standards of understanding research methods should be a high priority. Those psychiatrists interested in pursuing careers in research should consider attempting courses at the London School of Hygiene and Tropical Medicine or City University (and elsewhere) which provide an in depth exploration of issues central to the

methodology of behavioural science such as epidemiology, psychometrics and statistics.

The MSc in psychiatry at the Institute of Psychiatry directly assesses candidates' ability to referee a paper adequately as part of the end of year examination. It may be that journal editors should consider selecting referees by a similar small test – seeing whether their reviews adequately discriminate between papers whose faults and strengths are already known to editors.

### References

- ATKINSON M. (1994) Regulation of science by peer review. *Studies in the History and Philosophy of Science*, **25**, 147–158.
- BARBER, B. (1961) Resistance by scientists to scientific discovery. *Science*, **134**, 596–602.
- BORNSTEIN, R. F. (1990) Epistemic progress and journal page limitations: a proposal for increasing the base rate of manuscript acceptance in psychology journals. *American Psychologist*, **45**, 673–674.
- BRADLEY, J. V. (1982) Editorial overkill. *Bulletin of the Psychonomic Society*, **19**, 271–274.
- COTTON, P. (1993) Flaws documented, reforms debated at congress on journal peer review. *Journal of the American Medical Association*, **270**, 2775–2778.
- FINKE, R. A. (1990) Recommendations for contemporary editorial practices. *American Psychologist*, **45**, 669–670.
- ILTS, H. (1966) *Life of Mendel*. New York: Hafner.
- INGELFINGER F. J. (1974) Peer review in biomedical publication. *American Journal of Medicine*, **56**, 686–692.
- KUHN, T. S. (1970) *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- LOCK, S. (1986) Getting the balance right. *British Medical Journal*, **292**, 428–429.
- MARSH, H. W. & BALL, S. (1989) The peer review process used to evaluate manuscripts submitted to academic journals: interjudgmental reliability. *Journal of Experimental Education*, **57**, 151–169.
- O'CONNOR, M. (1978) *Editing Scientific Books and Journals*. Tunbridge Wells: Pitman-Medical.
- PETERS, D. P. & CECL, S. (1982) Peer review practices of psychological journals: the fate of published articles, submitted again. *Behavioral and Brain Sciences*, **5**, 187–195.
- WILKINSON, G. (1995) Jim Birley in conversation with Greg Wilkinson. *Psychiatric Bulletin*, **19**, 33–30.

Rajendra Persaud, *Consultant Psychiatrist, Maudsley and Bethlem NHS Trust, London SE5 8AZ*