

BLOWING THE WHISTLE ON REFEREES

DR PETER BOWBRICK¹

The research selectivity exercise presumes that the best papers get published in the best journals. The evidence suggests that on the contrary, the refereeing process actually discriminates against the best papers, and that many of the papers in the best journals are rubbish.

It has been argued for instance that if the refereeing system is 99% accurate, then 90% of what is published will be rubbish. If an Einstein writes a paper, there is a one per cent chance of it being wrongly rejected. There is also a one per cent chance of a really bad paper being accepted. Since there are 1000 bad papers for every one by an Einstein, 90% of what is published is bad. At 67% accuracy, there would be 500 bad papers published for every good one.

Unfortunately the refereeing system is nothing like 99% accurate. The evidence suggests that referees will usually agree that some papers are very bad indeed, and they will usually agree to reject (not accept!) a very good paper. For the rest, the judgement is random. For example, Ingelfinger examined the performance of referees for the New England Journal of Medicine. 500 papers were each refereed by two people. The two referees agreed only slightly more often than could be expected by chance. Since a quarter of the papers were considered bad by both reviewers, this suggests complete randomness on the papers that were not obviously bad. However, in fully 10% of cases one reviewer rated the paper "A" while another rated it "D". Similar studies on papers submitted to biomedical journals and social science journals have shown agreement only slightly better than random.

There is considerable evidence that the best papers are more likely to be rejected. Current Contents ran a series of articles by the authors of the most cited papers in the physical and biological sciences - those that were cited over 1000 times in ten years. Time after time the authors complained "I had more difficulty in getting this published than anything else I have written." Some of the more prolific authors in economics and statistics have found the same: it is easy to place a routine paper using or modifying standard techniques, but it is difficult to place an original, important or controversial paper. I know a case where one journal rejected a paper as rubbish, but another, of higher status, accepted it as being "the most important paper ever published in this journal".

The good papers are ones which present new theory or refute old theory, which give new ways of looking at things, or which present unexpected results. Referees who accept these papers must be willing to accept new ideas: they must accept both

¹Published in the Times Higher Education Supplement, Feb 10 1995
Copyright Peter Bowbrick, peter@bowbrick.eu 07772746759. The right of Peter Bowbrick to be identified as the Author of the Work has been asserted by him in accordance with the Copyright, Designs and Patents Act.

that they have been wrong in the past, and that they have missed what now appear to be obvious errors. This is a threat to their self esteem, and subconscious blocking mechanisms come into play.

Experiments have shown that referees are more likely to accept papers which support their own beliefs. Papers which had the identical method, analysis and discussion sections, but different results, were sent to referees by a psychology journal. They were much more likely to be accepted if the results supported the referee's theoretical position.

Similarly, papers with inconclusive or negative results are more often rejected.

There is also evidence that some journals have editors and referees drawn from half a dozen universities, who seem uncomfortable with the paradigms and analysis used elsewhere, and who seldom accept papers from other universities. The American Economic Review, the American Sociological Review and the Quarterly Journal of Economics have been accused of this at different times.

A paper is likely to be rejected if its methodological approach is not the same as the referees'. There are ivory towered theoreticians who will reject a paper as 'eclectic' if it has sufficient assumptions to approximate to reality, or as 'anecdotal' if it cites empirical data.

A paper which attacks and refutes the paradigm of the journal is almost certain to be rejected. I have several times had such papers violently rejected by the specialist journal, only to be accepted by more prestigious general journals. The referees did not publish a comment in reply.

At a more trivial level there is the referee who rejects a paper because his own work is not cited: I had a paper on a highly specialized area of the economics of quality criticized by a referee on the grounds that I had not read or cited his book on the philosophy of economics - a book that had not been published when the paper was written.

This unconscious and sometimes conscious bias had a big effect in the past, but the pressures are changing. The research selectivity exercise and the casualization of academic employment have introduced a new moral hazard. Every referee has a strong personal interest in whether or not a paper is published. If someone else publishes papers which challenge the paradigm, or publishes a string of good papers, the referee's career suffers. The referee may lose research income and research ratings as a result. This will result in lost promotion and possibly a lost job. The conflict between departments aiming to become research centres also has a strong influence on the referees future employment and job satisfaction. The ideal is to create a situation where a lot of individuals are each publishing a few mediocre papers in your paradigm, so you shine in a crowded mainstream paradigm.

The impact of the bias and the moral hazard is multiplied by the fact that some journals are competing to be regarded as top journals by having more referees than the competition, two, three or even more. The more referees there are, of course, the more likely it is that one of them will reject your paper on specious grounds.

A good journal is one that publishes good papers, not one that does not publish bad ones. Because of the weaknesses of the refereeing process and the Einstein effect, it must accept that it will publish some bad papers if it is going to be sure of publishing good ones. Some of the best journals, Nature or The Economic Journal under Keynes' editorship, achieve this by minimizing refereeing. The fact that a paper is published in an excellent journal is far from being a guarantee that it is good.

Most journals exclude the very bad and the very good and publish a random selection of the rest, with some preference given to papers with a common subject and method of presentation.

The research selectivity exercise commits the error a social scientist is warned against from the cradle: do not think that just because something can be measured, it is meaningful.

1108 WORDS

Dr Peter Bowbrick