Journal of Biological Response Modifiers
3:10–14 © 1984 Raven Press, New York

Point of View

Peer Review of Scientific Papers

J. Scott Armstrong

Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania

Suppose the following study were conducted:

One hundred previously published papers are selected from highly prestigious journals in various areas of medicine. Each has been published by a researcher at a prestigious school. The articles are disguised by using fictitious names for the authors and for the institutions. In addition, cosmetic changes (changes of title, revisions in abstracts and introductions, changes of tables to graphs and vice versa) are made to prevent the articles from being identified, but the content is not changed. After checking to ensure that there have been no major changes in editorial policy, each article is resubmitted to the same journal in which it was published within the previous 3 years. (Assume that these journals typically use two referees, and that they do not conceal the names of authors from referees.)

What predictions would you make for:

1. The percentage of the papers (of the 100 resubmitted) that would be detected as having been previously published?
2. The percentage of the papers not detected that would be rejected?
3. The percentage of the rejected papers that would be rejected because they did not add anything new?

This hypothetical experiment is an extension of the Kosinski study (1,2). In 1979, Ross typed Jerzy Kosinski's novel *Steps* (3) and submitted it, untitled, to 14 major publishing houses and 13 literary agents. This novel, which had won the prestigious National Book Award for fiction in 1969, was rejected by all, including Random House, the original publisher. No one recognized the book. (Ross later replicated his study, with similar results, using the script for the movie *Casablanca*.)

Unlike publishers of fiction, scientific journals operate with a highly developed peer review system. General standards are agreed on, and they are applied in an objective manner. After all, peer review provides the primary mechanism for awarding government research grants, prizes, and promotions. Peer review also provides the only practical way of allocating space in the leading scientific jour-

Received August 31, 1983; accepted October 31, 1983.

Address correspondence and reprint requests to Dr. Armstrong at Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104.

nals. Obviously, then, you could not generalize from the Kosinski study . . . or could you?

THE PETERS AND CECI STUDY

Peters and Ceci (4), inspired by the Kosinski study, resubmitted 12 papers to the same prestigious psychology journals that had published them within the previous 3 years. Fictitious names were used for the authors and institutions; the original authors had been from prestigious universities. (These journals did not use blind reviewing.) Cosmetic changes were made to disguise the papers, but the content was not changed. Here are the results:

1. Only three papers (25%) were detected as having been published previously.
2. Of the nine papers that went undetected, eight (89%) were rejected.
3. Of the rejected papers, not one was rejected because it added nothing new. In other words, none of the referees' reports said that the results were "old hat."

Are you surprised? Probably not. Slovic and Fischhoff (5) found that few scientists are surprised by research results . . . no matter what the results. When Slovic and Fischhoff dramatically reversed the results of a study and presented the two different versions to two equivalent groups of readers, neither group was surprised.

To assess surprise, I asked a convenience sample of 21 full professors from various fields to predict the results of the Peters and Ceci study (6). Their predictions for the three questions were: detected, 66%; rejected (of those reviewed), 42%; and nothing new (of those rejected), 46%.

These results can be contrasted with the actual results of 25, 89, and 0%, respectively. The differences for questions 1 and 2 are larger than those from the prediction of 50% (based on the assumption of no knowledge). In short, the results of the study by Peters and Ceci were surprising to a group of professors with much experience as authors, referees, and editors.

The article by Peters and Ceci was published along with open peer commentary (7). The contributors were drawn from a wide variety of disciplines (e.g., biological, physical, and social scientists). Comments were provided by grant administrators, reviewers, authors, advocates and critics of peer review, sociologists, and historians of science. I think the Peters and Ceci study makes a significant contribution to our knowledge about the publication of scientific work, in spite of some shortcomings. This conclusion was shared by a majority of the commentators.

Many people were not enthusiastic about the Peters and Ceci article. Earlier versions of it had been rejected by *Science* and by *American Psychologist*. Many of the editors for the 12 journals involved in the Peters and Ceci experiment were upset. Some of them withheld relevant information, making it more difficult to interpret the findings. One editor tried to sabotage the study by informing other editors (8). Threats were made against Peters and Ceci (9).

Apparently, Peters and Ceci were tampering with a religion. Those who point out shortcomings in this religion are dangerous people. Other examples are common. Mahoney (10) reported that his 1977 experiment on journal reviewing practices led to attempts to have him fired. Manwell and Baker (11) describe cases where plagiarists recycled work that had previously been published by others in different journals. Two academic "whistleblowers" who asked for an investigation of the plagiarism were dealt with harshly (one was fired), but little was done to investigate the alleged plagiarisms. Shortcomings in the review system are often tolerated, but pointing them out is not.

BIAS AGAINST NEW IDEAS

The bias against authors is disturbing. Possibly even more serious, however, is a bias against papers that oppose existing scientific beliefs. Galileo opposed existing beliefs in the 1600s, and, as a result, he suffered. Experimental studies show that this type of bias is still prevalent. For example, Mahoney (12) submitted a paper to 75 referees from the *Journal of Applied Behavior Analysis*. Some received results that confirmed the dominant hypothesis held by scientists in that field, while others received the identical paper except that the results were disconfirming. More referees rejected the study with disconfirming evidence than the one with confirming evidence. They rated the disconfirming study poorer on "relevance" and "methodology," despite the fact that these factors were identical in both versions. Additional evidence on referee bias is presented in the experiments by Abramowitz et al. (13) and Goodstein and Brazis (14). Apparently, referees have different standards for articles in which the results conflict with their own beliefs.

SCIENTIFIC INNOVATION

Major innovations, which tend to refute current wisdom, frequently come from outside the field. From the evidence above we would suspect that highly innovative articles would have difficulty gaining acceptance from major journals, particularly if they came from low-status sources and challenged commonly held ideas. It is not difficult to find examples of major scientific papers that were refused publication or were delayed for years by major journals (for descriptions of such cases, see refs. 15 and 16). Classic cases include Mendel's work in genetics and Mayer's discovery of the first law of thermodynamics.

POSSIBLE SOLUTIONS

Scientific journals provide the most important medium for communicating new ideas and discoveries. For example, Levitan (17), in a survey of 46 biomedical scientists, found that 85% ranked "the journal" as the most frequently used information source for keeping up to date. Given the limited space in journals, peer review is necessary. It is important, therefore, to decide how to improve peer review.

I suggest five procedures to improve peer review (although undoubtedly there are more possibilities). Some of these procedures are currently used by some journals.

1. *Blind reviewing*. Blind reviewing would reduce the prejudice against unknown authors from institutions perceived by the reviewer to be of low status.

Blind reviewing is inexpensive and should be favored unless further research proves otherwise. The only contrary evidence I have found is the pilot study by Perlman (18). That study found that for a given journal, papers published by authors from higher-status institutions were more highly cited than those published by authors from lower-status institutions.

2. *Referees nominated by authors*. Authors should be asked to suggest referees (this policy is used by *Science*) as well as who should not be used as referees. The authors are likely to know who are the experts in their area. More important, they may know which qualified referees would not reject their study because it conflicted with their own viewpoint or because they might hold prejudicial views on the authors or their research. Editors should try to select referees with different viewpoints on the problem. They should use the author's suggestions as one source (but not the only source) in selecting referees.

3. *Open peer review*. Many experts claim that referees are more open and objective when their identities are not revealed. Almost all journals keep the referee's identity secret. This secrecy is preferred by many referees: for example, Rowney and Zenisek (19), in their survey on reviewers of three journals for the Canadian Psychological Association, found that 56% were opposed to giving the reviewers' names with the published paper.

Another argument against open peer review is that the referee, in an attempt to avoid errors, would be overly cautious about being associated with a paper that reports unusual results. For example, Gordon (20) reported that medical editors rated the publication of incorrect papers as a more serious error than the rejection of a high-quality paper.

Other experts argue in favor of open peer review. They claim that referees would be motivated to do a more competent and fairer review if they had to defend their views to the authors, and if they were identified with the published articles (21).

A modified position would allow referees to decide whether they want their identities revealed for a given paper. My small-scale survey of professors found 45% willing to have their names revealed to the authors and to the readers (6). Another possibility is that referees could designate portions of their reviews to be signed and published.

The availability of referees' reviews would provide useful information to scientists. Few readers can devote the attention to a paper that is given to it by referees. Journals could publish the names and addresses of the referees for each paper so readers could write for copies of their review.

Referees who do an outstanding job on important papers could be offered an opportunity to publish their reviews. This policy might encourage referees to accept papers on important issues.

4. *Structured guides for reviewers*. One solution for situations in which evaluators use irrelevant criteria is to prepare a structured guide that contains the relevant criteria. Many journals use such guides. In my opinion the guidelines should emphasize aspects that are currently given little weight, such as the importance of the study and the objectivity of the methods. An example of such a guide has been published by Armstrong (22).

5. *"Note to Referees."* Authors of controversial papers could be encouraged to submit a "Note to Referees." This would describe the hypotheses, design of the study, methods, and data, but it would not include the results. Such a procedure would help the referee to make an objective evaluation of the importance of the study and the methodology. After responding to the "Note to Referees," the referee would open a sealed package containing the complete paper and proceed with the review. This procedure is currently being used by the *Journal of Forecasting*.

The above changes in peer review should improve the publication prospects for papers with significant contributions. This seems especially important in fields such as medicine in which advances are rapid.

REFERENCES

1. Ross C. Rejected. *New West* 1979;4:39–41.
2. Ross C. Rejecting published work: similar fate for fiction. *Behav Brain Sci* 1982;5:236.
3. Kosinski J. *Steps*. New York: Random House, 1968.
4. Peters DP, Ceci SJ. Peer-review practices of psychological journals: the fate of published articles, submitted again. *Behav Brain Sci* 1982;5:246–52.
5. Slovic P, Fischhoff B. On the psychology of experimental surprises. *J Exp Psychol* 1977;3:544–51.
6. Armstrong JS. Barriers to scientific contributions: the author's formula. *Behav Brain Sci* 1982;5:197–9.
7. Harnad S. *Peer commentary on peer review*. New York: Cambridge University Press, 1982.
8. Peters DP, Ceci SJ. Peer-review research: objections and obligations. *Behav Brain Sci* 1982;5:246–52.
9. Sieber JE. Whose ethics? Or the perils and ethics of studying powerful persons. *Soc Adv Soc Psychol Newslett* 1983;9(August):1–5.
10. Mahoney MJ. Publication, politics, and scientific progress. *Behav Brain Sci* 1982;5:220–1.
11. Manwell C, Baker CM. Reform peer review: the Peters and Ceci study in the context of other current studies of scientific evaluation. *Behav Brain Sci* 1982;5:221–5.
12. Mahoney MJ. Publication prejudices: an experimental study of confirmatory bias in the peer review system. *Cog Ther Res* 1977;1:161–75.
13. Abramowitz SI, Gomes B, Abramowitz CV. Publish or politic: referee bias in manuscript review. *J Appl Soc Psychol* 1975;5:187–200.
14. Goodstein LD, Brazis KL. Credibility of psychologists: an empirical study. *Psychol Rep* 1970;27:835–8.
15. Colman AM. Editorial role in author-referee disagreements. *Bull Br Psychol Soc* 1979;32:390–1.
16. Colman AM. Manuscript evaluation by journal referees and editors: randomness or bias? *Behav Brain Sci* 1982;5:205–6.
17. Levitan KB. Scientific societies and their journals: biomedical scientists assess the relationship. *Soc Stud Sci* 1979;9:393–400.
18. Perlman D. Reviewer "bias": do Peters and Ceci protest too much? *Behav Brain Sci* 1982;5:231–2.
19. Rowney JA, Zenisek TJ. Manuscript characteristics influencing reviewers' decision. *Can Psychol* 1980;21:17–21.
20. Gordon MD. The role of referees in scientific communication. In: Hartley J, ed. *The psychology of written communication*. London: Kogan Page, 1980:263–75.
21. Freese L. On changing some role relationships in the editorial review process. *Am Sociol* 1979;14:231–8.
22. Armstrong JS. Research on scientific journals: implications for editors and authors. *J Forecasting* 1982;1:83–104.