1998

Democracy Does Not Make Good Science: On Reforming Review Procedures for Management Science Journals

J. Scott Armstrong
University of Pennsylvania, armstrong@wharton.upenn.edu

Suggested Citation:
Armstrong, J. S. 1998. Democracy does not make good science: On reforming review procedures for management science journals. Interfaces 28(4) 88--91.

http://interfaces.journal.informs.org/

Hugh J. Miser (1998), "Reaction to Armstrong's 'Management folklore and management science,'" *Interfaces*, 28 (4), 81-93.

Reply by: J. Scott Armstrong, The Wharton School, University of Pennsylvania, "Democracy Does Not Make Good Science: On Reforming Review Procedures for Management Science Journals," *Interfaces*, 28 (4), 88-91.

I agree with almost everything that Miser says. The one major disagreement is that he believes changes in the review procedures would be of marginal importance whereas I believe that they are vital.

In his last sentence, he says, "There is important research on these questions [on editorial procedures] to be done." There is. On the other hand, much has already been done. I recently completed a review of research on journal peer review (Armstrong 1997). In that paper, I describe findings from 68 empirical studies, of which 12 were experimental or quasi-experimental. All but three studies were published since 1975. The quality of this research is high. So we might first examine what has been discovered from this research.

It is not possible for me to do justice to the accumulated evidence in this note. Some findings were counterintuitive to me at first, so they are likely to seem that way to other readers. However, my conclusions seem to reflect a growing consensus among people who do research on this topic.

One finding about peer review is reassuring. A given paper is generally better after review. This common belief is supported by blind ratings using independent panels and by surveys of authors. On occasion, however, authors make revisions that they believe are harmful to their papers in order to get them published.

Other findings are distressing. Journal review procedures are not effective in making decisions as to which in a set of papers should be published. This is illustrated by the experience of leading economists, as reported by Shepherd (1995). Typical review procedures have low reliability, use false cues, such as irrelevant statistical-significance tests, and are subject to biases. In particular, there is a bias against papers that present new and important findings. In other words, current journal-review procedures inhibit the flow of the most useful scientific information. This is not due to incompetent, mean-spirited editors. Inhibitions arise largely because of dedicated editors trying to follow institutional procedures. These procedures are intended to ensure fairness and quality, but the evidence suggests they have had little success I achieving these aims (Armstrong 1997).

Miser suggests that we can improve the low of scientific advances by appointing aggressive and outstanding researchers to be editors. I agree, and Franke, Edlund, and Oster (1990) provide evidence to support this position. Further research is needed.

In my opinion, the primary problem with current reviewing procedures is that reviewers are asked to recommend whether a paper should be published. Certainly voting seems to be fair and democratic. Editors typically claim that they use this information only as guidance in making their decisions. Despite editors like Miser and Peterson (Armstrong 1996), however, it is rare for editors to go against the vote of the reviewers. How do I know? It is reported in the research literature. Journal peer reviews are unreliable (for example, see Cicchetti 1991), yet editors closely follow the advice of those who recommend rejection. Kupfersmid and Wonderly's (1994, p. 56) summary of four empirical studies led them to conclude that papers receiving mixed reviews have a low probability of being accepted. In a study of 263 blind reviews for the *Journal of Counseling Psychology* from 1982 through 1983, Munley, Sharkin, and Gelso (1988) found that editors' decisions to publish were highly correlated with reviewers' recommendations. Marsh and Ball (1989) obtained similar results. Bakanic, McPhail, and Simon (1990) evaluated the manuscripts submitted to the *American Sociological Review* from 1977 to 1982 and concluded that "a single recommendation to reject often resulted in a rejection" (p. 378).

When authors appeal the vote by reviewers, editors typically side with the reviewers (Simon, Bakanic, and McPhail 1986). That is interesting considering that while authors spend hundreds of hours on their papers, referees average about four hours on their reviews (Jauch and Wall 1989; Kind, McDonald, and Roderer 1981; Lock and Smith, 1990; Yankauer, 1990). Also, reviewers are not required to follow scientifically prescribed procedures in their reviews. They typically rely solely upon their opinions. On the other hand, reviewers are often objective and they may have high expertise in some relevant areas.

It is a simple matter to encourage editors to seek important papers unconstrained by reviewers' votes: Reviewers should not be asked to make a recommendation as to publication. Why should this work? First, consider the respective motivations; journal editors are more likely to want to publish important findings, while reviewers want to prevent errors. Second, editors are typically selected from among the best researchers, whereas reviewers tend to be more typical of the average researcher. Third, editors are aware of the total set of competing papers and can thus make a more informed judgment about which papers are best, whereas reviewers see the paper in isolation. Finally, editors may be able to more effectively use the feedback if it is not accompanied by a numerical summary score. This latter issues was examined in a study by Slovic and MacPhillamy (1974). They found that a common measure across alternatives led decision makers to weight it heavily even when the believed dimension had validity. Cautioning subjects about the weight on a common irrelevant dimension did not remove its influence nor did feedback about correct weightings. Follow-up analyses of how subject made decisions indicated that they had no awareness of giving special emphasis to scores on a common dimension; they believed that they took each factor into proper consideration.

My recommendation to eliminate democracy in the reviewing process is one of a set of institutional reforms I proposed (Armstrong 1997). The purpose of the reforms is to help journals to focus on the communication of scientific innovations. That process has been subverted by the use of publication as a way to make personnel decisions. The advancement of science has taken a back seat to the advancement of scientists. This has led to procedures to ensure fairness in the reviewing process. I believe that scientific innovation will be best served by eliminating fairness as a criterion. Furthermore, I expect that the end result may be even more fair than it is now because it would be more likely to reward those who produce useful scientific innovations.

I am not aware of any direct tests of the procedure to request reviewers' feedback without recommendations. Of course, it has been used for invited papers, edited books, and special issues, apparently with good effects. The *International Journal of Forecasting* will be conducting a small trial on this.

Meanwhile, reviewers can act independently. Since completing my examination of the evidence on journal peer review. I have adopted a policy as a reviewer that I make no recommendations about publication. (Although it has not happened yet, I might make an exception to recommend the acceptance of a highly innovative paper.) I provide only comments about the content of the paper. The editors should decide, and some of them may conclude that their job is to publish innovative and useful research. To date, editors still seem happy to receive my reviews.

Acknowledgement

Fred Collopy, Richard Franke, Raymond Hubbard, and William T. Ross, Jr., provided comments on early drafts.

References

Armstrong, J. S. 1996, "The ombudsman: Management folklore and management science – On portfolio planning, escalation bias, and such," *Interfaces,* 46 (4), 28-42.

Armstrong, J. S. 1997, "Peer review for journals: Evidence on quality control, fairness, and innovation," *Science and Engineering Ethics*, 3 (1), 63-84.

Armstrong, J. S. 1982, "Barriers to scientific contributions: The author's formula," *Behavior and Brain Sciences*, 5 (2), 197-199.

Bakaric, V., McPail, C., and Simon, R. J. 1990, "If at first you don't succeed: Review procedures for revised and resubmitted manuscripts," *American Sociologist*, 21 (4), 373-391.

Cicchetti, D. V. 1991, "The reliability of peer review for manuscript and grant submission: A cross disciplinary investigation," *Behavioral and Brain Sciences*, 14 (March), 119-186.

Franke, R. H., Edlund, T. W., and Oster, F. 1990, "The development of strategic management: Journal quality and article impact,“ *Strategic Management Journal*, 11, 243-253.

Jauch, L. R. and Wall, J. L. 1989, "What they do when they get your manuscript: A survey of Academy of Management reviewer policies," *Academy of Management Journal*, 32, 157-173.

King, D. W., McDonald, D. D., and Roderer, N. K. 1981, *Scientific Journals in the United States: Their Production, Use, and Economics.* Stroudsburg, PA: Hutchison Ross.

Kupfersmid, J. and Wonderly, D. M. 1994, *An Author's Guide to Publishing Better Articles in Better Journals in the Behavioral Sciences*. Brandon, VT: Clinical Psychology Publishing Co.

Lock, S. and Smith, J. 1990, "What do peer reviewers do?" *Journal of the American Medical Association*, 263 (10), 1341-1343.

Marsh, H. W. and Ball, S. 1989, "The peer review process used to evaluate manuscripts submitted to academic journals: Interjudgmental reliability," *Journal of Experimental Education*, 57 (2), 151-169.

Munley, P. H., Sharkin, B., and Gelso, C. J. 1988, "Reviewer ratings and agreement on manuscripts reviewed for the *Journal of Counseling Psychology*," *Journal of Counseling Psychology*, 35 (2), 198-202.

Shepherd, G. B., ed., 1995, *Rejected: Leading Economists Ponder the Publication Process*, Sun Lakes, AZ: Thomas Horton and Daughters.

Simon, R., Bakanic, V., and McPhail, C. 1985, "Who complains to editors and what happens," *Sociological Inquiry*, 56, 259-271.

Slovic, P. and MacPhilllamy, D. 1974, "Dimensional commensurability and cue utilization in comparative judgment," *Organizational Behavior and Human Performance*, 11, 172-194.

Yankauer, A.1990, "Who are the peer reviewers and how much do they review?" *Journal of the American Medical Association*, 263, 1338-1340.