October 1982

Is Review by Peers as Fair as it Appears?

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

Postprint version. Published in *Interfaces*, Volume 12, Issue 5, October 1982, pages 62-74.
Publisher URL: http://www.aaai.org/AITopics/html/interfaces.html

Is Review by Peers as Fair as it Appears?

Abstract
Recent research shows that journal reviewing practices are neither objective nor fair. I propose a procedure to increase the likelihood of publishing important papers. This will be tested by *Interfaces* for a year.

Comments
Postprint version. Published in *Interfaces*, Volume 12, Issue 5, October 1982, pages 62-74.
Publisher URL: http://www.aaai.org/AITopics/html/interfaces.html

Published in *Interfaces*, 12 (October 1982), 62-74.

Is Review By Peers As Fair As It Appears?

Recent research shows that journal reviewing practices are neither objective nor fair. I propose a procedure to increase the likelihood of publishing important papers. This will be tested by *Interfaces* for a year.

Suppose the following study were conducted: One hundred previously published papers are selected from highly prestigious journals in various areas of management science. Each has been published by a researcher at a prestigious University.

The articles are disguised by using fictitious names for the authors and the institutions. In addition, cosmetic changes are made to prevent the article from being identified (change of title, revisions in abstract and introductions, changing tables to graphs and vice versa), but the content is not changed.

After checks to ensure that there have been no major changes in the editorial policy, each article is then resubmitted to *the same journal where it was published two or three years earlier.* Assume that these journals do not routinely conceal the names of authors from referees and that the journals typically use two referees. What predictions would you make for:

1) the percentage of the papers (out 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 paper that would be rejected because they did not add anything new?

Our hypothetical experiment is an extension of the Kosinski study [Ross 1979, 1982]. In 1979, Ross typed up Jerry Kosinski's novel *Steps* and submitted it, untitled, to fourteen major publishing houses and thirteen literary agents. This novel, which had won the prestigious National Book Award for fiction in 1969, was rejected by all, including Random House, its original publisher. No one recognized the book. Yet, when the results of Ross's study were published, not one of the publishers or agents seemed to feel as Ross did: that the review system for novels should be evaluated.

Unlike publishers of fiction, journals of science operate with a highly developed peer review system. Standards are agreed upon and applied in an objective manner, as well they should be. After all, peer review provides the primary mechanism for awarding prizes, research grants, promotion, . . . and fame. Peer review also provides the only practical way of allocating space in the leading academic journals in the management sciences. Obviously, then, you could not generalize from the Kosinski study. Or could you?

The Peters and Ceci Study

Peters and Ceci [1982a], inspired by the Kosinski study, resubmitted twelve papers to the same prestigious psychology journals that had published them within the previous three years. Fictitious names were used for the authors and institutions; the original authors had been from prestigious universities. (The 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 were undetected, eight (89%) were rejected.
3) Of the rejected papers, *none* was rejected because it added nothing new. In other words, none of the referees' reports contained anything saying that the results were old hat.

Are you surprised? Probably not; Slovic and Fischhoff [1977] found that few people are surprised by research results. Even 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 twenty-one full professors from various fields to predict the results of the Peters and Ceci (P&C) study [Armstrong 1982b]. Their predictions for the three questions were:

1) Detected	66%
2) Rejected (of those reviewed)	42%
3) Nothing new (of those rejected)	46%

The differences for questions 1 and 2 are larger than one could have obtained by the minimax prediction of 50% (based on the assumption of no knowledge)! In short, P&C's results were surprising to a group of professors with much experience as authors, referees, and editors.

The P&C paper was published in the June 1982 issue of the *Behavioral and Brain Sciences* along with open peer commentary from contributors drawn from a wide variety of disciplines: biological, physical, and social scientists, including editors, grant administrators, reviewers, authors, advocates and critics of peer review, sociologists, and historians of science. (it is also being published as a book by Cambridge University Press, *Peer Commentary on Peer Review, S.* Harnad, et al., planned 1982). I think the P&C study makes a significant contribution in spite of some shortcomings. My conclusion is shared by a majority of the commentators. It is an important study with dramatic findings. I hope it leads to further studies on the issue of peer review.

Many people were not enthusiastic about the P&C paper. Earlier versions of the paper had been rejected by *Science* and by the *American Psychologist.* Many of the editors for the twelve journals involved in the P&C 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 [Peters and Ceci 1982b]. Many of the commentators were also critical, particularly about the ethics of deceiving editors. (Experimental study often requires deception, but the matter seems especially serious when *editors* are deceived.)

Apparently, P&C are tampering with religion. Those who point out shortcomings in this religion are dangerous people. For example, Michael Mahoney [1982] reported that, his 1977 experiment on journal reviewing practices led to attempts to have him fired. Other examples are provided by Manwell and Baker [1982]. They describe cases where plagiarists recycled work that had been previously published in other journals. Two academic "whistleblowers" who spotted the plagiarism were dealt with harshly (one of them was fired). The plagiarists were not punished. Apparently shortcomings in the review system are tolerated, but pointing out shortcomings is not.

Bias against New Ideas

The bias against authors is disturbing, but possibly even more serious, is a bias against papers that oppose existing scientific beliefs. Galileo suffered from this bias. Experimental studies find this kind of bias to be still prevalent.

Mahoney [1977] 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. Apparently referees have different standards for articles where results conflict with their own beliefs.

Additional evidence on referee bias is presented in the experiments by Abramowitz, Gomes, and Abramowitz [1975) and by Goodstein and Brazis [1970].

Scientific Innovation

Studies of innovations show that they frequently come from outside their field. Major innovations tend to refute current wisdom. From the evidence above, we would suspect very innovative articles to have difficulty gaining acceptance from major journals, particularly if they came from low status sources and they challenged commonly held ideas. Indeed, it is not difficult to find examples of major scientific papers that were refused publication or were delayed for years by major journals [Colman 1979, 1982]. Classic cases include Gregor Mendel's work in genetics and Mayer's discovery of the first law of thermodynamics.

Is "peer review" simply a nice term for censorship?

A Proposal

Peer review is greatly relied upon in the management sciences where almost 80% of the papers are rejected by the prestigious journals. (In physics, on the other hand, almost 80% are accepted.)

Given the reward system in the social sciences and the current need for peer review, we should try to improve the peer review process. I suggest five procedures, (though 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 low status institutions. Blind reviewing is inexpensive and should be favored unless further research proves otherwise. (The only contrary evidence is found in the pilot study from Perlman, [1982]. That study suggests that for a given journal, papers published by authors from higher status institutions are actually better than those published by authors from lower status institutions.)

2) *Referees Nominated by Authors:* Authors should be asked to suggest referees as well as who should not be used as a referee. The authors are likely to know who the experts are in their area. More important, they may know which qualified referees would not reject their study because it conflicted with their own viewpoint. (This policy is used by *Science*.)

   Editors should try to select referees with different viewpoints on the problem. They would use the author's suggestions as one source in selecting referees.

3) *Open Peer Review:* Many experts claim that referees are more open and objective if their identity is not revealed. Almost all journals keep the referee's identity secret. This secrecy is preferred by many referees; for example, Rowney and Zenisek [1980], 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.

   Other experts argue in favor of open peer review. The proposed advantages are appealing. For example, referees should be more highly motivated to do a competent and fair review if they may have to defend their views to the authors and if they will be identified with the published papers. (Freese [1979] provides a detailed discussion on this issue.)

   A modified position would be to let referees decide whether they wanted their identity revealed for a given paper. My small-scale survey of full professors [Armstrong 1982b] found 45% willing to have their names revealed to the authors and to the readers. 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 can 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, used by the *Journal of Experimental Psychology: General*, would encourage referees to accept papers on important issues.

4) *Structured Guides for Reviewers:* The typical solution for situations where judges use false cues is to create a structured guide that contains the relevant cues. Many journals use guides. In my opinion, the guide should emphasize aspects that are currently given little weight such as the importance of the study and the objectivity of the methods used in the study. An example of such a guide is published in Armstrong [1982a].

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, anti 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 this "Note to Referees," the referee would open a sealed package containing the complete paper and proceed with the review.

I believe these to be useful proposals. In fact, I would like to use them for the next year as an editor of *Interfaces.* If you have *empirical research* on a topic relevant to the practice of management science, if you feel this research makes a major advance, and if you feel that the traditional review process has been unfair to your paper, I will review it for *Interfaces.* The requirements are that you must first exhaust the normal channels. Then

- Describe the controversial aspects of the paper and submit a reviewing history for the paper. The history would not be seen by the reviewers but would be included in the preface to the paper if it were published.
- Follow the procedures outlined above for the nomination of referees and for the "Note to Referees."
- Provide a $100 submission fee, to be used only for out of pocket expenses plus honoraria of $25.00 for each referee. Unused funds would be returned to *Interfaces* at the conclusion of this experiment.
- Be sure the paper is written in a manner that is comprehensible. If it is incomprehensible to me, I will return the paper along with your check.
- Follow the "Instructions to Authors" in this issue of *Interfaces.*

If convinced that the paper deals with an important topic, I will act as your advocate in trying to get the paper published. Although the Editor-in-Chief, Gary Lilien, will have veto power, he originally proposed the idea for the "Ombudsman" column and is interested in publishing important yet controversial work.

Peer review is not as fair as it appears. Nor is it as helpful to scientific achievements. I hope that the above procedures may help *someone*, *sometime*. While I wait, I will report to you on some of my own outrageous research.

Acknowledgments

Stevan Harnad and Baruch Fischhoff provided useful suggestions.

References

Abramowitz, Stephen I., Gomes, B., and Abramowitz, C. V. (1975), "Publish or politic: Referee bias in manuscript review," *Journal of Applied Social Psychology,* 5, 187-200.

Armstrong, J. Scott (1982a), "Research on scientific journals: Implications for editors and authors," *Journal of Forecasting,*1, 83-104.

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

Colman, Andrew M. (1979), "Editorial role in author-referee disagreements," *Bulletin of the British Psychological Society,* 32, 390-391.

Colman, Andrew M. (1982), "Manuscript evaluation by journal referees and editors: Randomness or bias?" *Behavioral and Brain Sciences,* 5, 205-206.

Freese, Lee (1979), "On changing some role relationships in the editorial review process," *American Sociologist,* 14, 231-238.

Goodstein, L. D. and Brazis, K. L. (1970), "Credibility of psychologists: An empirical study," *Psychological Reports*, 27, 835-838.

Kosinski, Jerzy (1968), *Steps.* New York: Random House.

Mahoney, Michael J. (1977), "Publication prejudices: An experimental study of confirmatory bias in the peer review system," *Cognitive Therapy and Research*, 1, 161-175.

Mahoney, Michael J. (1982), "Publication, politics, and scientific progress," *Behavioral and Brain Sciences*, 5, 220-221.

Manwell, Clyde and Baker, C. M. (1982), "Reform peer review: The Peters and Ceci study in the context of other current studies of scientific evaluation," *Behavioral and Brain Sciences*, 5, 221-225.

Perlman, Daniel (1982), "Reviewer 'bias': Do Peters and Ceci protest too much?" *Behavioral and Brain Sciences*, 5, 231-232.

Peters, Douglas P. and Ceci, Stephen 1. (1982a), "Peer-review practices of psychological journals: The fate of published articles, submitted again," *Behavioral and Brain Sciences*, 5, 187-195.

Peters, Douglas P. and Ceci, Stephen J. (1982b), "Peer-review research: Objections and obligations," *Behavioral and Brain Sciences*, 5, 246-252.

Ross, C. (1979), "Rejected," *New West*, 4, 39-41.

Ross, C. (1982), "Rejecting published work: Similar fate for fiction," *Behavioral and Brain Sciences*, 5, 236

Rowney, Julie A. and Zenisek, Thomas J. (1980), "Manuscript characteristics influencing reviewers' decisions," *Canadian Psychology*, 21, 17-21.

Slovic, Paul and Fischhoff, Baruch (1977), "On the psychology of experimental surprises," *Journal of Experimental Psychology: Human Perception and Performance*, 3, 544-551.

A comment from Martin Starr, Editor-in-Chief, *Management Science*

It has been suggested a number of times that *Management Science* mask authorship (called blind review). In each case, after study, the recommendation has been negative. The main reasons were: there was nothing conclusive about the extent of bias with open review, blind review promotes guessing about authorship using the reference list to make inferences about authorship, the guesses might also lead to bias, blind review prohibits checking other work (not cited) by the authors to determine possible duplication, blind review in which authors' citations are deleted is misleading, and there is additional editorial burden.

Thus censorship from the editors' and reviewers' point of view is difficult to justify unless it can be shown that significant bias exists with open review. That is what Scott Armstrong's article tries to do. However, we can account for the results he reports in another way. To me, the most obvious explanation is that the editors have made a bad choice of referees. I submit that this alternative hypothesis provides a reasonable fit with the study results. Good referees should have detected that the papers were previously published. Lacking that, they should at least have known the state of the art and, thereby, rejected the papers because they did not add anything new. Refereeing is an expert task, one part of which is knowing the state of the art and what has been published. Editors err if at least one referee is not fully aware in this way.

Does the editor ask the referees if they know the state of the art? Often, one is surprised to find that an "established" expert in methodology is less conversant with the published literature than a recently-graduated Ph.D.

The above is no criticism of blind review. It is critical of an argument for masking authorship based on a presumption of bias using studies which could produce the same results for other reasons that are, at least, as tenable.

Rather than treating the symptoms by using blind review, I suggest attacking the illness of inadequate refereeing through editorial means. Editors should know the strengths and weaknesses of their referees. They should augment their list by using suggestions from authors (a policy of *Management Science* for many years), and they should consult with referees about their appropriateness for each paper. Since open peer review (unmasking referees to authors) can possibly decrease the quantity of qualified referees, this step should be taken with care. I doubt that $25 paid to each referee can overcome the reluctance to be entirely frank with a colleague.

There's a lot to be learned about open peer review. What has happened to evaluations and recommendations where the candidate has access to the files? There's much to be learned about bias. In my experience, bias has little to do with author renown or school reputation, but is strongly affected by schools of thought. This includes concepts (e.g., the fuzzy set versus the axiomatic), and even mathematical notation forms the basis for affiliations. A good editor calls on referees who (when taken together) belong to all relevant schools, including those just forming, and then, even after revisions, may have to provide one or more umpires to make a go or no-go decision.

A reply to the comment by Martin K. Starr by J. Scott Armstrong

My proposal for blind review would not involve the deletion of the authors' citations. I am most concerned about the reduction of biases against authors who have no prior publications on the topic. Well-known authors can reveal their identity in the paper if they care to.

Honoraria for referees are not intended for the purpose of making the referee frank. Rather, they are provided as a token of appreciation. It is expected also that honoraria will lead to a faster reply from the referee.

A comment from Herbert F. Ayres, President of TIMS

I believe the effects Armstrong cites are quite real. I also believe none of them is really serious, except the bias against fundamental new work. That is very serious. I believe that bias to be fixable but not by polishing the referee process. That bias is a basic defect of structure which requires a new high level "check and balance" counterweight to correct it.

Because it is difficult to get qualified professionals to serve as departmental editors and referees, most of the time editors-in-chief are the captives of their departmental editors who are in turn prisoners of their referees. Editorial overrides are very rare for fear of losing people. As I shall argue below, I think that a Rejection Review Board reporting to the Editor-in-Chief and having referee override authority is strong enough medicine to cure the illness. I believe that weaker remedies at lower levels will fail.

My Background Biases

Much of what we publish in our journals is the exercise and extension of *Technology.* Little of it is the reporting of basic, important and new scientific research. There are few people who are really groundbreakers. They are scattered and hard to find. If the science and technology community had to depend on them for organization and maintenance, it would probably collapse; which brings me to that slippery and misleading misnomer, "Peer Review." Most referees are not groundbreakers. They represent the "establishment" and are in a conflict-of-interest position if they encourage the potential disruption of groundbreaking. Further, "Peer Review" is an impractical objective in the case of basic scientific research. If there are 50 persons among TIMS's and ORSA's 10,000 members whose contributions are still cited 50 years hence, that will be a lot. How do you sign them up as referees, since you don't know who they are yet? Indeed, if we published only what were really fundamental and new, the majority of scientific journals in the country would have to close down. It is absolutely OK that the journals contain material that is mostly exercising technology. Where else on earth could we write and read about it?

Who's a Referee?

Consider Professor Zee, an established full professor. He has tenure, is an authority in his field, does consulting, teaches, takes part in university affairs, is active in his professional associations, has been a departmental editor for a well-known journal, is working on his nth book, supervises dissertations and still does some of his own research. In the 31 years since he was a freshman, he has invested 9 solid man-years of work mastering his specialty. If its structure is basically disrupted, he will need up to a solid man-year to rework it. Where is he going to get a man-year? Nowhere, unless he turns his life inside out. As a referee of a paper that threatens to disrupt his life, he is in a conflict-of-interest position, pure and simple. Unless we're convinced that he, we, and all our friends who referee have integrity in the upper fifth percentile of those who have so far qualified for sainthood, it is beyond naive to believe that censorship does not occur. It need not be entirely conscious (shoddy methodology, sloppy reasoning, opaque exposition, etc.). If most referees lack the maturity of Professor Zee, many admire him as a role model and are "becoming him" just as fast as they can. Wouldn't most departmental editors have Zees as their referees if they could get them? Note that Zee is not a bad buy. Zee is a good guy; he's part of the very backbone of the science and technology community. He never said he was a saint; it is we who are asking him to be.

Comments Specific to Armstrong

Let's look at the Peters & Ceci results:

1) *Only 25% Recognition.* Most articles are of no earth-shaking import and, therefore, are hard to remember.

2) *89 % of the Rest Rejected.* Without the cues of well-known authors and institutions, the lack of earth tremors was more easily felt. If you don't believe this explanation, try rewriting and resubmitting the work of I. Newton, A. Einstein, J. C. Maxwell or J. W. Gibbs to a physics journal. Have a go with the basic work of R. Bellman, G. Dantzig, A. A. Markov or Kuhn and Tucker in our field.

Other Comments

1) Armstrong keeps saying "Peer Review," "Peer Review." *That's the problem.* We cannot get "peers" of real groundbreakers as referees: *Peer Review is* a label that papers over our basic problem.
   a. There are not enough groundbreakers to go around.
   b. It takes 10 or 20 years to identify one.
   c. Once you find one, what makes you think he'll be eager to be your referee?
2) The cited, unsaintly negative reactions of the editors and their friends were because they – entirely *correctly* – saw the checks on "Journal Reviewing Practices" as criticism of *themselves* and the structure of *their* system. They are responsible because they are in charge. Do you expect referees to be saintly while their oxen are gored when editors are not?

Armstrong's Proposal

1) *Blind Reviewing.* I'm all for it, but caveat editor – your rejection rate may rise.

2) *Referees Nominated by Authors.* I'm all for that too. The more referees, the weaker the chains on the departmental editors.

3) *Open Peer Review.* I'm very much against it beyond offering the option to the referee. Insist on it, and you'll drive away younger people who cannot stand up to the establishment without grave personal risk. You'll wind up with middle-aged fogeys like – you know – thee, Zee, and me!

4) *Structured Guide for Reviewers.* ROTS-O-RUCK. Just try to tell me how to review a paper. Go ahead, try! How many referees do you suppose share my pliable, open-minded attitude?

5) *Note to Referees. I* think authors should be permitted to say or not to say anything they want in a letter of transmittal to referees with no restrictions on timing or structure or content. I also believe the referee should be free to use any or none of it.

A Proposal for a Rejection Review Board

From what I've said above, it is clear that I believe bias against and censorship of important, fundamental and therefore disruptive new work is solidly inherent in the present structure of the journals of technology and science. While I enthusiastically applaud Armstrong's offer for advocacy and your idea of an "Ombudsman" column, I think more is required. If I were an Editor-in-Chief of a journal, this is what I would do.

First, I'd sell the idea of a Rejection Review Board to my departmental editors, emphasizing how it would help them with their referees. (I've heard that the shadow of the appeals court occasionally falls across the desk of a judge writing an opinion. Also, if there has to be an override, the heat is not on the Editors.)

Second, I'd try to get the departmental editors to pick the trio for the board and to help me sell those individuals on the idea. They should be established authorities with real clout. I am going to need real heavy artillery on the Board.

The Rejection Review Board should operate under rules like the following:

1) Any rejectee may petition the board for review by submitting (a) the disputed paper and all correspondence; (b) written acknowledgement that the Board's decision is final, and a written indemnification of all parties; and (c) a substantial fee to be returned only if the petition is successful.

2) The only claims which will lead to successful petitions are newness and importance. Attacks on accepted theory are explicitly encouraged.

3) If a rejecting referee wishes to have even informal interaction with the Review Board, he must do so under his own name. Behind the scenes communications from said referee to the Board should be viewed as the ethical equivalent of jury tampering.

4) The Board may consult anyone in its deliberations, including the writer and the 'referee if the latter is willing to drop his anonymity.

5) Majority vote of the Board determines, and a positive judgment commits the journal to publish.

6) The Board members serve the journal in no other capacity.

Third, with the Rejection Review Board in place, I would publicize its existence far and wide. If I could get away with it; I would publish all Board accepted papers as such, *calling attention to the controversy.* If I could swing it, I'd publish the disputed paper, the reasons for rejection, and the author's rebuttal all in the same issue.

It won't take much of this before many of the hot, controversial papers which can't make it through the system at other journals start showing up at ours. Other journals might get very cross at us. Wouldn't that be awful?

I am perfectly well aware that some of our editors have taken great personal pains to fight the bias against disruptive new work, and I hope these remarks will not make them angry at me. It's my view that under the existing organizational structure of most scientific journals, a diligent editor-in-chief can only reduce this bias. He needs something like a rejection review board to help him eliminate it.

A reply to the comment by Herbert F. Ayres by J. Scott Armstrong

I second the motion for a Rejection Review Board!

With respect to a structured guide for reviewers, our initial experience at the *Journal of Forecasting (JoF) is* that most referees chose to use it, and many have commented that it helped them to understand the aims and standards of the *JoF.* As a one-time industrial engineer, I am partial to the use of structured guides. I feel that I can do a better job as a referee when a journal provides me with a structured guide.

(Editorial Note: Dr. Ayres did not.have access to Dr. Starr's comment before writing his own. Comments from other readers are welcome.)

A comment from Stephen M. Robinson, Editor, *Mathematics of Operations Research*

I think the best way to start out the response is to review how I see Armstrong's paper. It seems to consist of (1) an account of an experiment involving parts of a novel; (2) an account of an experiment involving 12 papers in psychology; (3) some opinions by Mr. Armstrong, generally uncomplimentary to the present system of peer review; and (4) a proposal by Mr. Armstrong for modifying the peer review process. I have problems with the paper on two levels: first, its tone and organization; second, its concrete proposals.

First, I think the organization of the paper is rather deceptive. This paper is really just a statement by Mr. Armstrong that he doesn't think the present system is very good, and that he has a system he thinks is better. There is no reason for him to dress it up by describing experiments which are not visibly relevant to the field of management science. I personally can't see any reason why the Kosinski experiment, or the Peters and Ceci study in psychology, even if one accepts them as valid studies, should necessarily have any value as predictors of what will happen in operations research or management science. The fields are very different, and I see no reason to think that such results would carry over. Thus, it does not seem to me that the first half of Armstrong's paper has any valid purpose.

In addition to this organizational criticism, I have problems with the tone in which this paper is written. It is very arrogant, and it gives the impression that Mr. Armstrong has unlimited contempt for editors and others involved in the review and publication process. For example, he seems to interpret the reaction of editors to the Peters and Ceci study as being unjustifiable, and he comments that "experimental study often requires deception, but the matter is serious when *editors* are deceived." He does not seem to understand that experimentation involving human subjects also requires informed consent, and that people who find they have been made subjects of an experiment which has a cost to them (in this case, time) without their consent may quite justifiably be angry. This contemptuous tone is continued in comments, such as, "those who point out shortcomings in this religion are dangerous people," and "apparently shortcomings in the review system are tolerated, but pointing out shortcomings is not," and "the bias against authors is disturbing . . .

I think I have made it clear in the above that I have serious doubts about Mr. Armstrong's paper based upon the way it is organized and written. This is so because I have found in the past that people who write in this tone frequently do so because they are operating from a rather limited perspective and do not completely understand the system about which they are writing. Surely there are exceptions, but in this case I think Mr. Armstrong may not have looked carefully enough at the system he is criticizing. Indeed, there are some aspects that I do not think he understands. The first is that people are fallible, and that many of the decisions that he seems, to see as underhanded or as part of a conspiracy (e.g., rejection of innovative papers) may simply be cases of honest mistakes made by harried people. The other aspect, related to this, is that the refereeing system is not a system for selecting the best papers. It is a system for making the best selection possible with severly limited resources, particularly time, and with largely voluntary labor. There are many aspects of refereeing that one might change if one had a large staff of competent, paid referees who devoted their full time to reading and deciding on papers. We don't have this, and I don't think we can have it. When one realizes that the present system is a compromise between the ideal of absolutely objective, fully-informed opinion and the reality of very limited resources, then I do not think that the present review system does nearly as bad a job as Mr. Armstrong seems to think. In addition, the present system guards against some problems that could well arise if his suggestions (at the end of the paper) were followed. Some of these problems are:

1) Blind reviewing increases the time demands on referees, and given the fact that referees have limited time it will decrease the acceptability of refereeing to prospective referees. Many people will not review blind papers, taking the attitude that "I am doing this as a favor to the editor anyway, and why should I do a favor for someone who insults me by telling me that I cannot be expected to be objective?"

2) Having authors nominate referees is fine, but if editors are expected to use the nominees then this could aggravate the "buddy system" that already is claimed to exist in many fields.

3) If referees are required to reveal their identities to authors, then I believe we will get less honest and more evasive replies. People have to keep up professional and social relationships with others, and if giving an honest report would seriously damage a relationship that might be important to them, I suspect that many people would not give it. Thus, I think that open refereeing would decrease the quality of comment available to us.

To sum up what I have said, I do not think that this is a very good paper. I see no evidence that Mr. Armstrong understands very much about the way the reviewing process really works, and I think that he is jumping to conclusions about people's motives. Finally, I find the tone of the paper to be rather unpleasant, and although I am sure that the author is perfectly honest in trying to improve the system, I see little likelihood that his recommendations will do so.

Reply to the comment by Stephen M. Robinson by J. Scott Armstrong

Robinson's letter raises some key areas of uncertainty. Can we generalize to management science? Hopefully research on these issues will be done in management science. Until that time, I assume Robinson is advocating gradual change and low-cost strategies. I believe this to be good advice.

A comment from Seth Bonder, Past President of the Operations Research Society of America

I have no qualms about experimenting with any of his suggestions except perhaps the first (nomination of referees) which could likely lead to a "mutual publication society," wherein you accept mine and I will accept yours. A number of his other recommendations have been tried with variable success. It is my impression that quality referees are technically knowledgeable people in the field who are willing to devote the time to do a thorough review (approximately two to three days on a reasonably technical paper). Last thought: perhaps the reason physicists accept eighty percent of submitted papers is that eighty percent of them have substance worth publishing while only twenty percent of what we do is worth the archival space.