# Barriers to Scientific Contributions: The Author's Formula

J. Scott Armstrong
The Wharton School, University of Pennsylvania
Philadelphia, Pa. 19104

Published in *Behavioral and Brain Sciences*, 5 (June 1982), 197-199

Recently I completed a review of the empirical research on scientific journals (Armstrong 1982). This review provided evidence for an "author's formula," a set of rules that authors can use to increase the likelihood and speed of acceptance of their manuscripts. Authors should: (1) *not* pick an important problem, (2) *not* challenge existing beliefs, (3) *not* obtain surprising results, (4) *not* use simple methods, (5) *not* provide full disclosure, and (6) *not* write clearly. Peters & Ceci (P&C) are obviously ignorant of the author's formula. In their extension of the Kosinski study (Ross 1979; 1980), they broke most of the rules.

Why, then, is P&C's paper being published? In my search for an explanation, I learned the following from Peters: (a) After a long delay, the paper was rejected by *Science,* with advice that it would be appropriate for the *American Psychologist.* (b) After a long delay, the paper was rejected by the *American Psychologist.* This history illustrates the predictive power of the author's formula. Submission was meanwhile encouraged by the editor of the *Behavioral and Brain Sciences* – a journal specializing in peer interaction on controversial papers – and, after a final round of major revision, the paper was accepted for publication.

In this commentary, I describe how P&C violated many rules in the author's formula. It may be too late to salvage their careers, but the discussion should be instructive to other authors.

**Examined an important problem.** P&C examined whether the decision of prominent scientists to recommend a paper for publication constitutes evidence of that paper's scientific contribution. This strikes me as an important issue. It passed one test I use for importance: Would I discuss this issue with people outside my field? It has implications for the communication of scientific knowledge. Few researchers have dared to address it. Most of those who have done excellent work on this issue have met difficulties in getting their findings published in high-prestige journals. For example, Goodstein and Brazis (1970), Mahoney and Kimper (1976), and Mahoney (1977) were *not* published in journals with high prestige. Furthermore, Mahoney (1977) was rejected by *Science.*

**Challenged existing beliefs.** Scientists believe themselves to be competent and fair when they judge scientific contributions. An alternative hypothesis, such as, "The judgment by scientists of scientific contributions is seriously affected by irrelevant factors," is an affront to scientists. P&C reveal themselves to be insensitive to this fact. Their work does not try to make a "positive scientific contribution"; instead, it merely criticizes an existing belief. (P&C have obviously learned little from the well-publicized case of Galileo, namely, that some beliefs should not be examined.) P&C's use of the method of multiple hypotheses to examine existing beliefs was impressive. It was difficult to think of a reasonable alternative to current beliefs that they did not examine. The strategy of multiple hypotheses is unusual in the social sciences. Typically, the route to academic fame involves adopting a single dominant hypothesis and finding supporting evidence (Armstrong 1979; 1980a).

**Obtained surprising results.** To challenge existing beliefs is folly. To obtain evidence that these beliefs are wrong is sinful. (It is called the Second Sin by Szasz 1973.) P&C are sinners. They found that biases against the author or the author's institution play an important role in judgments of the value of a scientific contribution. These are surprising results.

For the journals used in P&C's study, the probability of an article being accepted was about 20%; the rate of acceptance for the papers resubmitted by P&C, 11%, was not significantly different. This suggests that we, as scientists, are merely engaged in a game of chance. Either P&C are wrong or we are fools!

Another way to interpret the results is to argue that there are identifiable reasons for acceptance or rejection but that these have nothing to do with the scientific contribution made by a paper. This reminds me of Webster's (1964) studies on the employment interview. Some factors did help explain who would be hired, but they had little to do with job performance. Instead, they related to the similarity between interviewer and interviewee. The decisions were usually made in the first five minutes (often in the first 30 seconds), and the interviewers were not aware of their decision-making process, although they claimed to have based it on the potential job performance of the interviewee. Perhaps the employment situation is analogous to the P&C study; that is, perceived similarity (e.g., in institutional background) is a measure of the author's competence. In fact, none of the reviewers had a background similar to the author's, since the rejected papers were from fictitious institutions.

I agree with P&C that bias against unknown authors and institutions provides the best explanation of their results. In fact, the evidence may be even stronger than they suggest. Noting that one of the manuscripts (manuscript "I" – see P&C's tables) was unanimously accepted (two referees plus two editors) while eight papers were unanimously rejected (16 referees plus 10 editors), I hypothesized that the author and institution for this sole accepted article may have *sounded* more authentic. (Hawkins, Ritter & Walter 1973 demonstrated that fictitious journals had high status among academics if they sounded theoretical.) Peters & Ceci (personal communication) provided some confirming evidence. The fictitious author of manuscript "I" had a common male name, Wade Johnston. This paper was originally submitted as having been from the University of North Dakota.

(In a convenience sample I conducted at the University of Pennsylvania's Wharton School, eight out of nine faculty members selected the University of North Dakota as the most prestigious from the list of six institutions used in the P&C study.) However, one other undetected paper had likewise been initially submitted with the institution identified as the University of North Dakota, and that paper was rejected.

You might not agree that P&C's results are surprising. This would be understandable. The experiment by Slovic and Fischhoff (1977) indicated that few scientists are surprised by the results of scientific studies, no matter what the results.

To determine the degree of "surprise" associated with P&C's results, I conducted a survey of five full professors at Drexel University and seventeen at the University of Pennsylvania. Most were from the social sciences, primarily management, but five were professors of physical sciences. (Four research assistants and I attempted to deliver personally a self-administered questionnaire to professors in these schools over a five-day period. Six questionnaires were administered by phone, Although there were few refusals, few professors were available; either they were out or occupied.) Fourteen of the 22 respondents (64%) reported that they were either editors or associate editors for one of the most prestigious scientific journals in their field.

The questionnaire contained a brief description of the P&C study. It was presented as a "proposed study" and the respondents were asked to predict what results would be obtained if the study were conducted with psychology journals. None of the respondents had previously heard about the P&C study. The results, summarized in Table 1, show sizeable differences between the predicted and actual outcomes. The respondents expected the journals to detect far more papers as having been previously published. Furthermore, they greatly underestimated the percentage of reviewed papers that would be rejected and greatly overestimated the percentage of the rejected papers for which the grounds for rejection would be that they added nothing new. If the respondents had assumed that they knew *nothing*, they could have minimized their maximum possible error for each question by predicting 50%. This prediction would have produced a smaller average error (38% rather than the 45%). Applying the survey outcome to the P&C study, one finds that only five papers would have been expected to go undetected through the review process and two to have been rejected. Only *one* paper would have been rejected for reasons other than that it "added nothing new" (vs. the eight such papers in the P&C study). In short, P&C's results were very surprising, The surprise was much higher among the social scientists than among the physical scientists. (Additional details on the survey are available from this commentator.)

**Table 1. Predicted Versus Actual Outcomes of P&C Study**

| | Average predicted % from survey (n = 22) | Actual % (P&C) |
|---|---|---|
| Papers detected | 66 | 25 |
| Papers rejected (out of total number reviewed) | 42 | 89 |
| Papers rejected as adding "nothing new" (out of total number rejected) | 46 | 0 |

**Used simple methods.** P&C examined alternative hypotheses by using simple methods. These methods seemed appropriate, and it was difficult to see a need for greater complexity. In short, P&C lacked the methodological rigor desired for publication in a prestigious journal. They would benefit from studying Siegfried (1970), who showed how even tbe simplest ideas ean be made complex: He provided a rigorous proof that 1 + 1 = 2.

**Failed to provide full disclosure.** To obtain information from some journals, P&C had to promise confidentiality (personal communication with Peters). It is interesting that scientific journals find it necessary to suppress relevant information. Does this protect science – or does it merely protect scientists?

P&C were unable to provide full disclosure. This omission represents the only time that their study clearly followed the author's formula. This shortcoming is unfortunate. How can we be sure that P&C did the study? Hoaxes are not unknown in science. Mahoney (1979, notes 91-95) provides a short bibliography of scientific hoaxes and deceptions, and recent deceptions are described in Trafford (1981). Or perhaps P&C made errors in their analyses: Wolin's (1962) study showed errors to be common for published studies in psychology.

Access to the original data would be helpful in evaluating P&C. What papers were used in the study? Who were the referees? What did the referees say in their reports?

**Wrote clearly.** Many items in the author's formula can be overlooked if only authors would write obtusely. Obtuse writing also seems to yield higher prestige for the author (Armstrong 1980b).

P&C almost followed the rule. Their Gunning fog index was about 18.[1] This represents material appropriate for a second-year graduate student and is typical for scientific papers. (For example, I calculated an average Gunning fog index of 18.7 for papers published in 58 sociology journals.) P&C's note 1 was certainly an excellent example of fog.

**Recommendations.** On the basis of prior research and their own study, P&C offer suggestions for improving the review system in journals. Three of their suggestions are of particular importance: structured guides for referees, open peer review, and blind refereeing.

> **Structured guides for referees.** A structured guide can clarify the aims of the journal, which should reduce the likelihood of bias due to irrelevant factors such as similarity in beliefs or backgrounds. I suggest that articles be reviewed according to a structured guide designed to refute the author's formula. That is, positive ratings should be sought for importance, challenges to current beliefs, surprising results, simple methods, full disclosure, and clear writing.[2]
>
> **Open peer review.** Some observers claim that referees will be more open in their criticism if their identity is kept secret from the authors and readers. Anonymity is certainly the traditional practice. In my survey of faculty members, 12 of the 22 respondents (55%) said they, as referees, would object to having their names revealed to the authors. Also, 12 others objected to having their names published along with the papers they reviewed.
>
> Nevertheless, I agree with P&C that an open reviewing system would be preferable. It would be more equitable and more efficient. Knowing that they would have to defend their views before their peers should provide referees with the motivation to do a good job. Also, as a side benefit, referees would be recognized for the work they had done (at least for those papers tbat were published).
>
> Open peer review would also improve communication. Referees and authors could discuss difficult issues to find ways to improve a paper, rather than dismissing it. Frequently, the review itself provides useful information. Should not these contributions be shared? Interested readers should have access to the reviews of the published papers.

---

[1] The Gunning fog index was calculated in the following way. $G = 0.4(S + W)$, where S is the average sentence length and W is the percentage of words with three or more syllables (not including prefixes or suffixes).

[2] We have attempted to do this in the editorial manual for our *Journal of Forecasting*. Copies are available from the author.

For important issues, referees could publish their review along with the paper. The format would be similar to that eurrently used by the *Behavioral and Brain Sciences*. Such a procedure would provide a favorable bias for the acceptance of papers dealing with important issues. [See editorial note following this commentary. Ed.]

**Blind refereeing.** My conclusion, based on the prior research cited by P&C, is that blind refereeing shou1d be used by journals. Some studies have found it to be helpful (Schaeffer 1970). Although other studies have indicated no need for blind refereeing (P&C might also have included the experiment by Mahoney, Kazdin & Kenigsberg 1978, and the survey by Kerr, Tolliver & Petree 1977), none has shown blind refereeing to be harmful. Furthermore, the cost of blind refereeing is negligible.

In the survey described above, my respondents thought that the reviewers would be able to guess the author and institution for about one-third of the papers. Furthermore, in P&C's study only three of the 38 reviewers (8%) were able to detect papers that had already been published in leading journals by individuals from prestigious institutions. In short, blind refereeing would be expected to be relevant for most papers.

Most journals do not use blind refereeing (e.g., Coe & Weinstock 1967 found that only 26% of the major economics journals used blind reviewing). Judged in light of the earlier pattern of research results, P&C provide eompelling evidence in favor of blind refereeing. Finally, 14 of the 18 professors expressing an opinion in my survey (78%) thougbt that journals should use blind refereeing, so this change should be an easy one to make.

**Fate of the author's formula.** Academicians do not believe that the author's formela exists. In my survey, the 22 professors were asked the extent to which they thought each factor increased (+2) or decreased (-2) the likelihood of publication in "the highest prestige journals in your field." Table 2 summarizes the results. Respondents felt that authors should pick important problems, obtain surprising results, and write clearly. The only agreement, and it was modest, was that simple methods should be avoided.

**Table 2. Academics' Opinions of the Author's Formula**
(-2 = greatly decreases to +2 = greatly increases)

| | **Effect on acceptance** |
|---|---|
| If the author does | |
| not pick an important problem | -1.6 |
| not challenge existing beliefs among scientists | -0.2 |
| not provide surprising results | -1.1 |
| not use simple methods | +0.4 |
| not provide full disclosure | -0.7 |
| not write clearly | -1.1 |

In contrast to these academics, I believe that the empirical evidence supports the existence of the author's formula (Armstrong 1982). The P&C case represents only a portion of the surprising research on scientific journals. We should examine this evidence and then take steps to penalize, rather than reward, use of the author's formula.

**Further research.** I wish that I had done the P&C study. One hopes that their study will be replicated and extended by others in fields other than psychology. The possibility of replications of this study should improve the vigilance of editors and referees: perhaps the next paper they receive has already been published. Particularly important would be a replication with journals that provide blind reviews. This will help to determine the relative importance of two of the most likely hypotheses in P&C: Is it a game of chance or is it bias against unknown authors and institutions?

---

Author's Note: This commentary was published along with the lead article by Peters & Ceci (1982).

Editor's Note:. *BBS* does not have a policy of open peer review (see C. Belshaw's commentary, this issue). Submitted manuscripts receive multiple, anonymous review (although referee anonymity is optional). Open peer commentary is accorded only to accepted articles (and then, of course, the referees are among those who are invited to submit a commentary). The *Journal of Experimental Psychology: General* has a policy of occasionally copublishing referee reports with accepted articles, and *Speculations in Science and Technology* sometimes publishes dissenting reviewers' exchanges of correspondence with authors (see W. M. Honig's commentary, this issue). R. A. Gordon's commentary discusses a similar proposal. Ed.

## References


Armstrong, J.S. (1980a), "Advocacy as a scientific strategy: The mitroff myth," *Academy of Management Review*, 5, 509-11.

Armstrong, J.S. (1980b), "Unintelligible management research and academic prestige," Interfaces, 10, 80-86.

Armstrong, J.S. (1979), "Advocacy and objectivity in science," *Management Science* 25, 423-28.

Coe, R. K. and I. Weinstock (1967), "Editorial policies of major economic journals," *Quarterly Review of Economics and Business*, 7, 37-43.

Goodstein, L. D. and K. L. Brazis (1970), "Credibility of psychologists: An empirical study," *Psychological Reports*, 27, 835-38.

Hawkins, R. G., L. S. Ritter and I. Walter (1973), "What economists think of their journals," *Journal of Political Economy*, 81, 1017-32.

Kerr, S., J. Tolliver and D. Petree (1977), "Manuscript characteristics which influence acceptance for management and social science journals," *Academy of Management Journal*, 20, 132-41.

Mahoney, M. J. (1979), "Psychology of the scientist: An evaluative review," *Social Studies of Science*, 9, 349-75.

Mahoney, M. J. (1977), "Publication prejudices: An experimental study of confirmatory bias in the peer review system," *Cognitive Therapy and Research*, 1, 161-75.

Mahoney, M. J. , A. E. Kazdin and M. Kenigsberg (1978), "Getting published: The effects of self-citation and institutional affiliation," *Cognitive Therapy and Research*, 2, 69-70.

Mahoney, M. J. and T. P. Kimper (1976), "From ethics to logic: A survey of scientists," *Scientist as Subject*, M. J. Mahoney, ed. 187-93. Cambridge, MA:Ballinger.

Peters, D. P. and S. J. Ceci (1982), "Peer-review practices of psychological journals: The fate of published articles, submitted again," *The Behavioral and Brain Sciences*, 5, 2, 187-195.

Siegfried, J. J. (1970), "A first lesson in econometrics," *Journal of Political Economy*, 78, 1378-79.

Szasz, T. (1973), *The Second Sin*, London: Routledge and Kegan Paul.

Trafford, A. (1981), “Behind the scandals in science labs,” *U. S. News and World Report*, March 2, p. 54.

Webster, E. C. (1964), *Decision Making in the Employment Interview.* Montreal: Eagle.

Wolin, L. (1962), “Responsibility for raw data,” *American Psychologist*, 17, 657-58.