

Perspectives on Psychological Science

<http://pps.sagepub.com/>

Masked Reviews Are Not Fairer Reviews

Nora S. Newcombe and Mark E. Bouton

Perspectives on Psychological Science 2009 4: 62

DOI: 10.1111/j.1745-6924.2009.01102.x

The online version of this article can be found at:

<http://pps.sagepub.com/content/4/1/62>

Published by:



<http://www.sagepublications.com>

On behalf of:



[Association For Psychological Science](http://www.sagepub.com/content/4/1/62)

Additional services and information for *Perspectives on Psychological Science* can be found at:

Email Alerts: <http://pps.sagepub.com/cgi/alerts>

Subscriptions: <http://pps.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Issues in Publishing, Editing, and Reviewing

Masked Reviews Are Not Fairer Reviews

Nora S. Newcombe¹ and Mark E. Bouton²¹Temple University and ²University of Vermont

ABSTRACT—*Many people believe that reviewers are less likely to produce unfair and biased reviews when they do not know the identity of the authors. However, there is surprisingly little evidence of such bias, there is little evidence that masked review is effective in addressing bias, and masked review has potential costs and drawbacks to the research community. More empirical work on the existence and nature of bias is needed.*

Journal articles began as scientific communications among a group of scientists who knew each other, writing letters to the Royal Society and the like. However, as science has expanded and grown more concerned with objective evaluation, many disciplines have evolved traditions in which reviewers do not know the identities of aspiring authors and/or these authors do not know the identities of reviewers. This commentary concentrates on the masking of author identity to reviewers. We mention whether or not reviewers should sign their reviews only briefly.

Disciplines differ as to whether both kinds of masking or only one or the other are used, as well as in whether masking is optional or required. Masking is less common in the physical and medical sciences than in the social sciences (Yankauer, 1991; see also Blank, 1991) and was endorsed by only 39% of the membership in a survey of the Federation of American Societies for Experimental Biology (Gidez, 1991). A survey of 96 medicine-related journals revealed that only 18.6% use masked review (Cleary & Alexander, 1988).

The common assumption of proponents is that when the identities of authors are masked and reviewers concentrate only on the research itself, evaluation is more objective. In this commentary, we reflect on whether this assumption is valid. To anticipate, we argue that masked review is a practice that exists to combat a fact not yet clearly in evidence (bias in reviewers);

that it may not work to combat such bias, if any exists; and that it has significant costs.

IS THERE EVIDENCE OF BIAS WHEN REVIEWS ARE NOT MASKED?

There are indeed some reasons to think that there might be reviewer bias when the identity of authors is known. Some kinds of evidence are indirect (e.g., studies of unconscious prejudice, or analog studies of hiring using matched resumes with male vs. female names; this literature is reviewed by Valian, 1999). There is also more directly relevant evidence. Peters and Ceci (1982) and Ceci and Peters (1984) randomly assigned more or less prestigious affiliations to real manuscripts (submitted again as new papers) and found that those listing lower prestige institutions received more negative reviewer evaluations. A study by Wenneras and Wold (1997) reported gender bias in records of evaluations of actual fellowship applications in Sweden. A recent meta-analysis of gender bias in evaluations of grants found a small but significant bias against women (Bornmann, Mutz, & Daniel, 2007).

However, there is also reason to question the argument that unmasked review results in significant bias. First, although somewhat indirectly relevant to bias, there is evidence that masked and unmasked reviews are of equivalent quality (Justice, Cho, Winkler, Berlin, & Rennie, 1998; van Rooyen, Godlee, Evans, Smith, & Black, 1998). If there were bias, one might expect that unmasked reviews might be rated as lower in quality. Second, there are studies that evaluate bias and find none. Gilbert, Williams, and Lundberg (1994) found that unmasked review did not result in gender bias in the *Journal of the American Medical Association*. Marsh, Jayasinghe, and Bond (2008) report no gender bias in reviews of Australian grant proposals, including no effect of whether the reviewer's gender matched or did not match the gender of the applicant. A RAND report found little bias in evaluation of federal grant applications in the U.S., with two possible exceptions (http://rand.org/pubs/research_briefs/RB9147/RAND_RB9147.pdf). Garfunkel, Ulshen, Hamrick, and Lawson (1994) did not find that institu-

Address correspondence to Nora S. Newcombe, Department of Psychology, Temple University, Weiss Hall, 1701 North 13th Street, Philadelphia, PA 19122; e-mail: newcombe@temple.edu.

tional prestige correlated with positivity of reviews for major papers submitted to the *Journal of Pediatrics*, although there was a correlation for brief reports. In commenting on the Bornmann et al. meta-analysis, Marsh et al. (2008) pointed out that bias in the studies reviewed showed a wide range, from 22% in favor of men to 23% in favor of women, suggesting perhaps that more research is necessary on what accounts for this range. Third, and most important, there are data that not only directly challenge the assumption of bias, they even suggest that masked review can backfire and have consequences opposite those intended by proponents. In a study of the prestigious *American Economic Review*, Blank (1991) found not only that there was no evidence of gender prejudice in unmasked review, but also that masked review actually favored authors from high-prestige institutions.

IS THERE EVIDENCE THAT ANY BIAS IN UNMASKED REVIEWS ARE REDRESSED BY MASKED REVIEWS?

For the sake of argument, however, let us concede that there may be reviewer bias when authorship is known that relates to issues such as gender, ethnicity, seniority, or institutional prestige. The next logical question is whether masking author identity is an effective solution. A key problem is whether masking is successful. Although Ceci and Peters (1984) found that reviewers could guess author identity for 26%–36% of manuscripts, a figure that is substantial but arguably not disabling to the efficacy of masking, other data suggest that masking is unsuccessful in even higher proportions of cases (Blank, 1991; Cho et al., 1998; Fischer, Friedman, & Strauss, 1994; Justice et al., 1998; van Rooyen et al., 1998; Yankauer, 1991). High rates of guessing authorship may be especially found when no effort is made to limit self-referencing, which is very common and very hard to redact without affecting the flow of an argument (McNutt, Evans, Fletcher, & Fletcher, 1990; Moosy & Moosy, 1985). In some fields, educated guesses about authorship can easily be based on apparatus sections, descriptions of subjects or participants, etc. Even more disturbingly, the data (McNutt et al., 1990; Moosy & Moosy, 1985) indicate that reviewers with more reviewing experience, more publications, and more time in research (i.e., highly qualified reviewers) were especially likely to be able to judge author identity successfully.

An additional issue is that masked review almost never involves the editor not knowing the identity of the authors. As Franzini (1987) noted, doubts about bias may persist as long as editors know authors' identities. There is disturbing evidence relevant to this point in an examination of articles submitted to the *Personality and Social Psychology Bulletin*. High-prestige authors and male authors received more favorable editorial decisions from both male and female editors, even after controlling for reviewer evaluations (Petty, Fleming, & Fabigarr, 1999). Although this argument might suggest that masking should extend to editors, such a system would be cumbersome to

administer and would have significant costs of both a monetary and an intellectual nature. In addition, recall that other studies have not shown gender bias even in journals not using masked review (Blank, 1991; Gilbert et al., 1994), which suggests that there is much more to understand about whether and when there is bias and how to eliminate it.

There seems to be only a handful of experiments that actually compared the fates of randomly selected manuscripts subjected to masked versus unmasked review. In her study of manuscripts submitted to *The American Economic Review*, Blank (1991) found that masked manuscripts were significantly less likely to be accepted for publication, with the trend being significant for male-authored papers but not female-authored papers. The authors suggested that female-authored manuscripts were not as affected by masking because “. . . more women are in institutional settings that are less negatively affected by blind refereeing, particularly lower-ranked universities and colleges” (p. 1055). (Masked review did not influence acceptance rates of papers at the highest ranking and lowest ranking institutions.) In another study, van Rooyen et al. (1998) asked editors to evaluate the quality of the reviews written about masked and unmasked manuscripts submitted to the *British Medical Journal*. They found no difference in the quality of reviews of masked and unmasked manuscripts, and the two types of manuscripts did not differ in the final decision to publish. Fisher et al. (1994) also found no difference in the recommendations given by masked and unmasked reviewers of manuscripts submitted to the *Journal of Developmental Pediatrics*. Justice et al. (1998) asked editors and authors to rate the quality of masked and unmasked reviews of manuscripts submitted to five medical journals. There was no difference in the ratings made by either editors or authors. In partial contrast, McNutt et al. (1990) studied manuscripts submitted to a single journal (*Journal of General Internal Medicine*) and found that editors rated the quality of blind reviews to be higher than that of nonblind reviews. However, blind and nonblind reviewers did not differ in their recommendations for publication, and, interestingly, authors did not rate the quality of blind reviews as different from nonblind reviews. Finally, Godlee, Gale, and Martyn (1998) sent masked and unmasked versions of a paper that had been doctored to include errors. Blind and nonblind reviewers did not differ in the number of errors they detected, although blind reviewers were less likely to recommend rejection. In total, the results of experiments studying the effects of masking manuscripts suggest surprisingly little benefit of implementing a masked review policy.

ARE THERE COSTS TO MASKED REVIEW?

The argument might still be made that, even if masked review does not actually remove bias (if it exists), masked review is nevertheless perceived as fairer and should be adopted for that reason. The argument would be persuasive, however, only if the costs associated with masking are minimal. We believe that

there are, in fact, several costs potentially associated with required masked review. First, if reviewers do not know seniority, they may provide less tutorial review for inexperienced authors, which is more commonly although not invariably provided when authors are known to be junior. Second, masked review may lead to a general reduction in openness and civility in science. In fact, many critics of the current peer review system (e.g., Epstein, 1995; Surwillo, 1986; see this issue) have called for reviewers to sign their reviews (open review—the second kind of masking issue we discussed in opening) as a remedy for what are seen as injustices and even cruelties in review. Third, masked review reinforces the (possibly incorrect) perception that there is gender-based or ethnic-based bias in review process.

In summary, masked review is a method that is designed to address a problem (bias) for which there is surprisingly little evidence. To date, there is little evidence that masked review is effective in redressing bias, even if we concede that such bias exists. And masked review has potential costs and drawbacks to the research community. More empirical work on the existence and nature of bias is needed, including how various systems other than masked review might work to combat it (see Marsh et al., 2008, for suggestions).

REFERENCES

- Blank, R.M. (1991). The effects of double-blind versus single-blind reviewing: Experimental evidence from *The American Economic Review*. *American Economic Review*, 81, 1041–1067.
- Bornmann, L., Mutz, R., & Daniel, H.D. (2007). Gender differences in grant peer review: A meta-analysis. *Journal of Infometrics*, 1, 226–238.
- Ceci, S.J., & Peters, D. (1984). How blind is blind review? *American Psychologist*, 39, 1491–1494.
- Cho, M.K., Justice, A.C., Winker, M.A., Berlin, J.A., Waeckerle, J.F., Callahan, M.L., & Rennie, D. (1998). Masking author identity in peer review: What factors influence masking success? *Journal of the American Medical Association*, 280, 243–245.
- Cleary, J.D., & Alexander, B. (1988). Blind versus nonblind review: Survey of selected medical journals. *Drug Intelligence and Clinical Pharmacy*, 22, 601–602.
- Epstein, S. (1995). What can be done to improve the journal review process. *American Psychologist*, 50, 883–885.
- Fisher, M., Friedman, S.B., & Strauss, B. (1994). The effects of blinding on acceptance of research papers by peer review. *Journal of the American Medical Association*, 272, 143–146.
- Franzini, L.R. (1987). Editors are not blind. *American Psychologist*, 42, 104.
- Garfunkel, J.M., Ulshen, M.H., Hamrick, H.J., & Lawson, E.E. (1994). Effect of institutional prestige on reviewers' recommendations and editorial decisions. *Journal of the American Medical Association*, 272, 137–138.
- Gidez, L.I. (1991). The peer review process: Strengths and weaknesses. *Serials Librarian*, 19, 75–85.
- Gilbert, J.R., Williams, E.S., & Lundberg, G.D. (1994). Is there gender bias in JAMA's review process? *Journal of the American Medical Association*, 272, 139–142.
- Godlee, F., Gale, C.R., & Martyn, C.N. (1998). Effect of the quality of peer review of blinding reviewers and asking them to sign their reports. *Journal of the American Medical Association*, 280, 237–240.
- Justice, A.C., Cho, M.K., Winkler, M.A., Berlin, J.A., & Rennie, D. (1998). Does masking author identity improve peer review quality? *Journal of the American Medical Association*, 280, 240–242.
- Marsh, H.W., Jayasinghe, U.W., & Bond, N.W. (2008). Improving the peer-review process for grant applications: Reliability, validity, bias, and generalizability. *American Psychologist*, 63, 160–168.
- McNutt, M.D., Evans, R.T., Fletcher, R.H., & Fletcher, S.W. (1990). The effects of blinding on the quality of peer review. *Journal of the American Medical Association*, 263, 1371–1376.
- Moosy, J., & Moosy, Y.R. (1985). Anonymous authors, anonymous referees: An editorial exploration. *Journal of Neuropathology and Experimental Neurology*, 263, 225–228.
- Peters, D., & Ceci, S.J. (1982). Peer review practices of psychological journals: The fate of published articles submitted again. *Behavioral and Brain Sciences*, 5, 187–196.
- Petty, R.E., Fleming, M.A., & Fabrigar, L.R. (1999). The review process at PSPB: Correlates of interreviewer agreement and manuscript acceptance. *Personality and Social Psychology Bulletin*, 25, 188–203.
- Surwillo, W. (1986). Anonymous reviewing and the peer-review process. *American Psychologist*, 41, 218.
- Valian, V. (1999). *Why so slow? The advancement of women*. Cambridge, MA: MIT Press.
- van Rooyen, S., Godlee, F., Evans, S., Smith, R., & Black, N. (1998). Effect of blinding and unmasking on the quality of peer review: A randomized trial. *Journal of the American Medical Association*, 280, 234–237.
- Wenneras, C., & Wold, A. (1997). Nepotism and sexism in peer review. *Nature*, 387, 341–343.
- Yankauer, A. (1991). How blind is blind review? *American Journal of Public Health*, 81, 843–845.