Efficacy of Double-Blind Peer Review in an Imaging Subspecialty Journal

E.E. O'Connor, M. Cousar, J.A. Lentini, M. Castillo, K. Halm and T.A. Zeffiro

AJNR Am J Neuroradiol published online 17 November 2016
http://www.ajnr.org/content/early/2016/11/17/ajnr.A5017
Efficacy of Double-Blind Peer Review in an Imaging Subspecialty Journal

E.E. O’Connor, M. Cousar, J.A. Lentini, M. Castillo, K. Halm, and T.A. Zeffiro

ABSTRACT

BACKGROUND AND PURPOSE: Many scientific journals use double-blind peer review to minimize potential reviewer bias concerning publication recommendations. However, because neuroradiology is a relatively small subspecialty, this process may be limited by prior knowledge of the authors’ work or associated institutions. We sought to investigate the efficacy of reviewer blinding and determine the impact that unblinding may have on manuscript acceptance.

MATERIALS AND METHODS: For manuscripts submitted to the American Journal of Neuroradiology (AJNR) from January through June 2015, reviewers completed a brief anonymous questionnaire after submitting their evaluations, assessing whether they were familiar with the research or had knowledge of the authors or institutions from which the work originated.

RESULTS: The response rate for 1079 questionnaires was 98.8%; 12.9% of reviewers knew or suspected that they knew authors, and 15.3% knew or suspected that they knew the associated institutions. Reviewers correctly identified the authors in 90.3% of cases and correctly stated the institutions in 86.8% of cases. Unblinding resulted from self-citation in 34.1% for both authorship and institutions. The acceptance rate when reviewers knew or suspected that they knew the authors was 57/137 (41.6%) and 262/929 (28.2%) when reviewers did not. The acceptance rate when reviewers knew or suspected that they knew the institutions was 60/163 (36.8%) and 259/903 (28.7%) when they did not. The Fisher exact test showed that author (P < .038) and institution (P < .039) familiarity was associated with greater manuscript acceptance.

CONCLUSIONS: While the AJNR process of double-blind peer review minimizes reviewer bias, perceived knowledge of the author and institution is associated with a higher rate of manuscript acceptance.
in a manuscript assigned to them or knowledge of the manuscript authors. Peer reviewers’ knowledge of author and/or institution identity may positively or negatively influence the content, quality, and final review recommendations. Thus, we sought to investigate the frequency with which perceived unblinding occurs and whether unblinding has an effect on manuscript acceptance. We hypothesized that perceived identification of the authors of a manuscript or related institutions would be associated with higher acceptance rates and that geographic residence of the reviewers or authors could independently affect the probability of manuscript acceptance.

MATERIALS AND METHODS

All manuscript reviewers submitting evaluations for AJNR from January 2015 to June 2015 were asked to answer a brief questionnaire using the electronic Manuscript Central system interface (Table 1). To evaluate the efficacy of blinding, the reviewers’ questionnaire replicated items from a study by Jagsi et al.21 Reviewers were asked whether they thought they could identify any of the principal authors or institutions from which the work originated. Possible responses for both of these questions included, “I know I could,” “I suspect I could,” or “I do not think I could.” Reviewers who thought or suspected that they knew the principal author of the manuscript and/or the institution from which the work originated, 138 (86.8%) were correct. Of the 159 reviewers who stated that they knew the institutions from which the work originated, 138 (86.8%) were correct.

To assess potential geographic influences on manuscript disposition, we recorded ZIP/country codes of the corresponding authors and reviewers and coded them by using the following regions: Northeast United States, Southeast United States, Midwestern United States, Mountain States United States, West Coast United States, North America (other than the United States), South America, Europe, Asia excluding China, Africa, and Australia. The regions of the corresponding authors and reviewers were treated as independent variables to investigate the effects of geography on the acceptance rate.

### RESULTS

The response rate for the 1079 questionnaires offered to reviewers was 98.8%. Of the 1066 responses, 137 (12.9%) knew or suspected that they knew the principal author of the manuscript and 163 (15.3%) reviewers knew or suspected that they knew the institution from which the work originated. Of the 154 reviewers suspecting that they had knowledge of the manuscript authors, 139 (90.3%) correctly identified the authors of the submitted manuscripts. Of the 159 reviewers who stated that they knew the institution from which the work originated, 138 (86.8%) were correct.

From the pool of reviewers who provided explanations for their suspected knowledge of authors and institutions of origin, unblinding resulted from self-citation in 34.1% of occurrences for both authorship and institutions.

The rate of acceptance when reviewers knew or suspected that they knew the authors was 57/137 (41.6%), and the rate of acceptance when reviewers did not identify the authors was 262/929 (28.2%). With a 2-sided Fisher exact test for count data, the null hypothesis that there was no association between perceived unblinding of the author and the rate of acceptance of the manuscript was rejected ($P = .038$; odds ratio $= 1.45$; 95% confidence interval, 1.016–2.059) (Table 2 and Fig 1A). An odds ratio of 1.45 measures the higher odds of the manuscript being accepted if perceived author unblinding occurs.

The rate of acceptance when reviewers knew or suspected that they knew the institutions from which the work originated was...
A, B. Mosaic plots graphically displaying data from 2 × 2 contingency tables. A, Unblinding of the author is associated with a higher rate of manuscript acceptance. B, Unblinding of the institution is associated with a higher rate of manuscript acceptance. There are 4 cells, and the area of each cell represents the frequency of each of the 4 unique combinations of variable levels. The ‘standardized residual’ is the residual divided by its SD. Therefore, the standardized residual rating represents the degree to which the 2 categorical variables are independent of each other, in units of SD. Colors represent the deviation in each cell from those expected from the null hypothesis that the 2 variables are not associated, with the color intensity indicating the degree of “surprise” associated with rejection of the null hypothesis (http://www.r-tutor.com/elementary-statistics/simple-linear-regression/standardized-residual).

Table 3: Contingency table showing manuscript counts and percentages for the relationship between manuscript disposition and author or institution unblinding in row percentages

<table>
<thead>
<tr>
<th>Manuscript Accepted</th>
<th>No</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author unblinded</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No 635 (71.4%)</td>
<td>254 (28.6%)</td>
<td>100%</td>
</tr>
<tr>
<td>Yes 112 (63.3%)</td>
<td>65 (36.7%)</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Institution unblinded</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No 634 (71.4%)</td>
<td>254 (28.6%)</td>
<td>100%</td>
</tr>
<tr>
<td>Yes 113 (63.5%)</td>
<td>65 (36.5%)</td>
<td>100%</td>
</tr>
</tbody>
</table>

*Both author and institution unblinding are associated with a greater chance of manuscript acceptance.

With a 2-sided Fisher exact test, the null hypothesis that there was no association between perceived unblinding of the institution and the manuscript acceptance rate was rejected ($P = .039$; odds ratio = 1.44; 95% confidence interval, 1.006–2.037) (Table 3 and Fig 1B). With a 2-sided Fisher exact test, the null hypothesis that there was no association between perceived author unblinding and perceived institution unblinding was rejected ($P < .0001$, odds ratio = 6920).

To examine whether the manuscript acceptance rate was dependent on reviewers guessing correctly, we used a 2-sided Fisher exact test with the null hypothesis that there was no association between correctly guessing and manuscript acceptance. The resulting odds ratio was 0.46 ($P = .18$).

With a 2-sided Fisher exact test, the null hypothesis that there was no association between the corresponding author’s geographic location and manuscript acceptance was rejected. Corresponding authors from the Southeast United States, Midwest United States, and West Coast United States were more likely to have a manuscript accepted, and corresponding authors from the Midwest United States were less likely to have a manuscript rejected ($P < .0001$). Authors from Asia and China were less likely to have manuscripts accepted and were more likely to have manuscripts rejected (Fig 2A).

With a 2-sided Fisher exact test, the null hypothesis that there was no association between reviewer geographic location and manuscript acceptance was rejected. Reviewers from the Midwest were more likely to accept a manuscript, and reviewers from Asia were less likely to accept and more likely to reject a manuscript ($P = .0035$) (Fig 2B).

The rate of AJNR manuscript acceptance during the 6-month study period was 29.9%. This acceptance rate was consistent with journal acceptance rates during a longer period, including 2013 (21.3%), 2014 (23.6%), and 2015 (28.2%). Across the United States, we observed strong geographic variation in reviewer and author density (Fig 3). Regional density was similar for authors and reviewers, with a strong association between the locations of authors and reviewers. These patterns appear to be related to the regional density of diagnostic radiologists reported by the Centers for Medicare and Medicaid Services (https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/medicare-provider-charge-data/physician-and-other-supplier.html).

DISCUSSION

Scientific progress depends on both the dissemination of ideas and their rigorous critique. Despite concerns about bias in the peer review process, most researchers still believe that double-blind peer review is essential for the appraisal of claims of new knowledge and effective communication with the scientific community, motivating the research community to grapple with the question of how to improve the existing review processes. In a single-blind review process, the reviewers’ identities are concealed from the author, but author identity is revealed to reviewers. While this is a commonly used model for peer review in the scientific community, double-blind peer review is perceived as more objective and fair. Its perceived benefits include protection of authors from potential reviewer bias and enabling reviewers to be forthright in their manuscript critiques without fear of interpersonal or political conflicts. To test these assumptions, we examined the incidence of perceived manuscript unblinding, examining the relationship between reviewers’ perceived knowledge of principal authors and institutions of manuscripts submitted to AJNR and their likelihood of acceptance.

Although the incidence of perceived unblinding in our study was relatively low compared with findings from studies evaluating its incidence in other journals, when perceived author or institution unblinding occurred, there was higher likelihood of manuscript acceptance.

A number of possible mechanisms could explain our results. Reviewers could feel compelled to accept the work of a known and respected investigator. Alternatively, it may be difficult to blind
reviewers to authors with longer research histories and more publications or who may have a history of submitting higher quality manuscripts. Dissociating the effects of manuscript quality from author and institution name recognition could be difficult without amassing more observations than were available for this study because we would need to accrue enough instances of high manuscript quality and low author or institution recognition to allow accurate estimates of these effects. An explanation of the international author effects may be that investigators whose first language is not English or who have not spent time in Western countries may not have written English skills commensurate with the challenges inherent in scientific writing, leading to lower perceived quality of the associated work.

While AJNR maintains a policy of author anonymity, we found that authors either forgot or intentionally failed to eliminate mention of institutions, affiliations, or previous work in 34.1% of instances of perceived unblinding. Other studies have also cited self-referencing patterns as a frequent source of unblinding.24 It may be possible to increase authors’ compliance with the policy on author anonymity by having journal staff remove self-citations in the manuscript before sending it out for review. Instructions for authors should also emphasize avoiding the use of the first person in their writing. For example, phrases such as “we have shown that” should be replaced with “Jones et al have shown that.”25 It has been recommended that authors avoid the use of 3 or more references from the same individual and avoid citations that are “in press,” because these typically reflect work by the same author or institution.25 In the latter case, reviewers may not have access to these unpublished references. These tasks might increase the workload for journal administrative staff. However, such additional effort may decrease the authors’ vulnerability to potential reviewer bias and may also improve the quality of the reviews.26,27

In some instances of unblinding, a reviewer may have previously reviewed the manuscript as a submission to a different journal before AJNR received it because there is a small pool of experts available for certain specialized topics. In a study evaluating factors that influence successful blinding, it was found that in 3 medical journals with a long-standing policy of blinding author identity, reviewers who spent more time in the United States, with the exception of the northeast and mountain states, manuscripts are accepted at a higher rate. Manuscripts from Asia and China have a lower-than-expected acceptance rate. Reviewers from the Midwest are more likely to accept a manuscript than would be expected, and reviewers from Asia are more likely to reject manuscripts. The area of each cell represents the frequency of each of the unique combinations of variable levels. The “standardized residual” is the residual divided by its SD. Therefore, the standardized residual rating represents the degree to which the 2 categoric variables are independent of each other, in units of SD (http://www.r-tutor.com/elementary-statistics/simple-linear-regression/standardized-residual).

FIG 2. Mosaic plots graphically displaying data from 12 × 2 contingency tables. A, Author location is associated with a differential rate of manuscript acceptance. In the United States, with the exception of the northeast and mountain states, manuscripts are accepted at a higher rate. Manuscripts from Asia and China have a lower-than-expected acceptance rate. B, Reviewer location is associated with a differential rate of manuscript acceptance. Reviewers from the Midwest are more likely to accept a manuscript than would be expected, and reviewers from Asia are more likely to reject manuscripts. The area of each cell represents the frequency of each of the unique combinations of variable levels. The “standardized residual” is the residual divided by its SD. Therefore, the standardized residual rating represents the degree to which the 2 categoric variables are independent of each other, in units of SD (http://www.r-tutor.com/elementary-statistics/simple-linear-regression/standardized-residual).

FIG 3. Reviewer locations for accepted and rejected manuscripts, while highly associated, are not distributed uniformly across the United States. This density distribution follows the distribution of diagnostic radiologists.
research were less likely to be successfully blinded.29 While using reviewers with less research and reviewing experience increased the success of blinding by journals, it is uncertain how this strategy affects review quality. One study showed that younger age, lower academic rank, employment at a strong academic institution, knowledge of the editor, and previous referee experience are reviewer characteristics associated with higher quality reviews.27,29 Higher review quality has also been shown to be associated with reviewers with training in epidemiology and statistics.30 Another study has shown that younger and more experienced referees tend to provide stricter assessments of manuscript quality than more senior or less experienced colleagues.31 It has also been shown that authors and editors perceived no difference in the quality of blinded and unblinded reviews.16,20,32

Peer review is intended to improve the publication process by helping editors select high-quality articles that are appropriate for the journal. It is also intended to improve the written presentation of the articles selected for publication.33 Double-blind peer review is designed to protect both authors and reviewers. Our study shows that the double-blind peer review process used at AJNR effectively maintains double anonymity in the most instances.

At a time when the millennial generation is joining the medical and scientific community, the possibility of an open or hybrid review system could be considered.34 In an open system, scientists post their articles on dedicated Web sites to undergo a review process in which authors and reviewers are known to each other.22,35 After nominated referees and other interested scientists post their comments regarding the work, the disposition for publication is determined. Proponents of this system argue that it increases transparency, incentivizes reviewers to be constructive in their critiques, and expedites the publication process.22,35 The British Journal of Psychiatry conducted a randomized controlled trial of open peer review, in which referees were asked to sign their reviews. Not only did most referees agree to sign their reviews, but the signed review quality was higher and more courteous than anonymized reviews. Signed reviews also took longer to complete.36

When the British Medical Journal conducted a randomized trial of referees revealing their identity to authors in the peer review process, there were no effects on editors’ ratings of the quality of the reviews provided or on manuscript disposition or the time taken to review. There was, however, a likelihood of referees declining to review.32 Some journals that have tried an open review system have had less favorable results. For example, when Nature tried an open peer review system in 2006, only a small proportion of authors opted to participate in the open system, and of authors who did participate, few received technically substantive comments.37 Currently, most journals are reluctant to adopt such a system because it can increase the time to final disposition, result in several article versions that may be confusing, may be associated with reviews of lesser quality than those from invited reviewers, and have fewer valuable reviews.13

A hybrid review system in which only selected articles undergo the open peer review process may provide the scientific community with exposure to this process before making a determination about its general viability. Nevertheless, whether an open or hybrid review system would be better for AJNR than the current double-blind peer review system is an open question because the current system is largely effective in minimizing potential biases against authors.

An editor takes responsibility for the quality of advice on the disposition of manuscripts. A judicious editor can prevent many of the abuses that a more open review system is intended to prevent, such as hostile comments, unsubstantiated criticisms, and delay of competitors’ manuscripts.38 We believe that scientific evaluation of the peer review process is of value to scientific journals because it will support them in maintaining the credibility of the publication process and demonstrating commitment to improving the process when needed. In addition, government agencies have taken interest in the peer review process during the past decade, emphasizing the importance of maintaining confidentiality,39 improving the integrity of the prepublication process, and encouraging research groups to optimize their review systems.34 Perhaps there could be more emphasis in the AJNR editorial process on the authors’ responsibility to maintain anonymity. For example, instead of stating “our previous work,” authors could be encouraged to say that “previous studies have shown.”

**Study Limitations**

In this study, reviewers’ experience with the manuscript blinding process was studied during a 6-month period. Sampling the review process for a longer time and allowing assessment of changing reviewer pool effects are needed for generalization of our findings. Also, some geographic regions had relatively few associated authors or reviewers; this feature limits the strength of the inferences that can be drawn from these data. Our study was observational rather than experimental in design, due to the need for AJNR to uphold its policy of anonymity to authors and reviewers. While a randomized experimental design would allow equal numbers of blinded and unblinded reviews, reviewers’ knowledge of whether they had been placed in the unblinded or blinded groups could bias their assessment of the manuscripts and influence the acceptance rate in ways that would be difficult to determine. Highly cited or well-published authors are more likely to submit high-quality manuscripts with an associated high likelihood of acceptance, and manuscripts from these authors are more likely to be recognized. Our study did not attempt to assess the confounding effects of manuscript quality, a potentially important factor influencing manuscript acceptance. Although it is possible that perceived unblinding and manuscript quality both influenced acceptance rates, further studies will be needed to isolate these effects. Finally, although our results are derived from sampling manuscripts in the relatively small field of neuroradiology, larger studies may allow generalization to a broader range of biomedical research.

**CONCLUSIONS**

The process of double-blind peer review used by AJNR is largely effective in minimizing reviewer bias. However, perceived unblinding of authors or institutions is associated with a higher rate of manuscript acceptance. There is also an association between author or reviewer geographic location and manuscript acceptance. Our results should motivate further study of double-blind peer review with a larger sample for a longer period.
REFERENCES

1. Glantz SA, Bero LA. Inappropriate and appropriate selection of ‘peers’ in grant review. JAMA 1994;272:114–16 CrossRef Medline
23. Ware M. Peer review in scholarly journals: perspective of the scholarly community—results from an international study. Information Services and Use 2008;28:109–12
24. Katz DS, Proto AV, Olmsted WW. Incidence and nature of unblinding by authors: our experience at two radiology journals with double-blinded peer review policies. AJR Am J Roentgenol 2002;179:1415–17 CrossRef Medline
25. Krinsky G. How to avoid “unblinding” the peer review process. AJR Am J Roentgenol 1999;172:1474 CrossRef Medline
35. Cartledge E. Peer review steps out of the shadows. Physics World 2007;20:29 CrossRef