Masking Author Identity in Peer Review

What Factors Influence Masking Success?

Mildred K. Cho, PhD; Amy C. Justice, MD, PhD; Margaret A. Winker, MD; Jesse A. Berlin, ScD; Joseph F. Waeckerle, MD; Michael L. Callaham, MD; Drummond Rennie, MD; and the PEER Investigators

Context.—In a previous study, we found that masking success was higher at a journal that masked reviewers to author identity. We hypothesized that masking policy or other factors could be associated with masking success.

Objectives.—To evaluate differences in success of masking reviewers to author identity at 7 biomedical journals and to identify factors associated with these differences.

Design.—Written questionnaire.

Participants.—Reviewers at 3 journals with a long-standing policy of masking author identity (Annals of Emergency Medicine, Epidemiology, and Journal of the American Geriatrics Society) and 4 journals without a policy of masking author identity (Annals of Internal Medicine, JAMA, Obstetrics & Gynecology, and Ophthalmology).

Main Outcome Measures.—Masking success (percentage of reviewers successfully masked) and reviewer characteristics associated with masking.

Results.—There was no significant difference in masking success between journals with a policy of masking (60%) and those without (58%) (P = .92). We found no association between masking success and a policy of masking when adjusted for the reviewer characteristics of age, sex, years of reviewing experience, number of articles published, number of articles reviewed, percentage of time spent in research, editorial experience, or academic rank (odds ratio [OR], 1.3; 95% confidence interval [CI], 0.64-2.8; P = .43). In a multivariable analysis of reviewer characteristics, reviewers spending a greater percentage of time in research, the only significant predictor of masking success, were less likely to be successfully masked (OR, 1.01; 95% CI, 1.00-1.02) (P = .04).

Conclusions.—Masking success appears unrelated to a journal policy of masking, but is associated with reviewers’ research experience and could be affected by other characteristics. Using reviewers with less research and reviewing experience might increase masking success, but the effect on review quality is unknown.

From the Center for Bioethics, University of Pennsylvania, Philadelphia (Dr Cho); Division of General Internal Medicine and Program for Health Care Research, Department of Veterans Affairs Medical Center, Cleveland, Ohio, and University Hospitals of Cleveland, Case Western Reserve University (Dr Justice); Division of Biostatistics, Department of Biostatistics and Epidemiology and Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, Philadelphia (Dr Berlin); Department of Emergency Medicine, University of Missouri/Kansas City School of Medicine (Dr Waeckerle); Division of Emergency Medicine, University of California, San Francisco School of Medicine (Dr Callaham); Institute for Health Policy Studies, University of California, San Francisco (Dr Rennie). Dr Winker is Senior Editor and Dr Rennie is Deputy Editor (West), JAMA.

A complete list of the members of the Peer review Effectiveness Evaluation Research (PEER) Investigators appears at the end of this article.

Presented at the Third International Congress on Peer Review in Biomedical Publication, Prague, Czech Republic, September 18, 1997.

Corresponding author: Mildred K. Cho, PhD, Center for Bioethics, University of Pennsylvania, 3401 Market St, Suite 320, Philadelphia, PA 19104-3308 (e-mail: mcho@mail.med.upenn.edu).

MASKING peer reviewers to author identity has been suggested as a means to improve quality and fairness of reviews. In a previous study, we found that masking reviewers to author identity in the peer review process was not uniformly successful at different journals, and that incomplete masking success limited our ability to assess the effect of masking on review quality. In particular, we noted that the Annals of Emergency Medicine had significantly higher rates of masking success than the other journals. Because the Annals of Emergency Medicine was the only journal in our study that had a long-standing policy of masking, we hypothesized that their policy might be responsible for the high rate, either because reviewers were accustomed to not knowing author identity or because the way in which the manuscript was prepared differed from other journals. Differences in characteristics of specialties might also be responsible.

In this study, we explored two questions: (1) is masking success associated with having a long-standing policy of masking? and (2) is masking success associated with reviewer characteristics?

METHODS

Journals and Procedure

In this study we included 7 journals, 3 with a long-standing policy of masking author identity (Annals of Emergency Medicine, Epidemiology, and Journal of the American Geriatrics Society) and 4 journals that did not mask author identity (Annals of Internal Medicine, JAMA, Obstetrics & Gynecology, and Ophthalmology).

Between August and December 1996, each journal enrolled manuscripts that met the following inclusion criteria: (1) the manuscript was sent for external peer review, (2) the manuscript reported original research, including meta-analyses but excluding case reports or letters, and (3) the authors did not object to having their manuscripts enrolled. Each journal masked eligible manuscripts by removing author and institutional identity from the title page, running headers or footers, and acknowledgments of manuscripts. Self-references in the text were not removed. Each journal then sent the manuscript to at least 2 reviewers along with a 1-page questionnaire. The reviewers were asked to return the manuscript and review to the journal and to return the questionnaire to the study coordinator in a separate, preaddressed envelope.

Each journal enrolled at least 20 manuscripts. Approximately one quarter of these were randomized to the usual reviewing process and were not analyzed in this study if the usual process was not to mask reviewers. Data from Annals of Internal Medicine, JAMA, Obstetrics & Gynecology, and Ophthalmology and data from 22 of 42 total manuscripts at Annals of Emer-
We used \( \chi^2 \) tests for analysis of categorical variables, and \( t \) tests for analysis of continuous variables. We used unconditional logistic regression models with single variables to estimate the unadjusted odds ratio (OR) for the outcome variable—masking success—associated with each of the potential predictor variables. Subsequently, we fit a multivariable logistic regression model including all the potential predictors to estimate the ORs for each variable, adjusted for all other variables in the model. Non-normally distributed continuous variables were also analyzed after log transformation, but transformation did not substantially change the results or conclusions. Results of untransformed data are reported here. For all tests of significance, we used 2-tailed \( \alpha = .05 \).

### RESULTS

#### Reviewer Response Rate and Masking Success Rate

The overall response rate of reviewers was 87% (287/328). No eligible reviewers declined to participate in the study. Overall, 60% of reviewers were successfully masked (95% confidence interval [CI], 54%-65%). The response rates and masking success rates by journal are shown in Table 1.

Table 1.—Success of Masking Reviewers by Masking Policy and by Journal

<table>
<thead>
<tr>
<th>Journal</th>
<th>Masked Manuscripts Enrolled, No.</th>
<th>Unadjusted OR (95% CI)†</th>
<th>Adjusted OR (95% CI)‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without masking policy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annals of Internal Medicine</td>
<td>20</td>
<td>90 (18/20)</td>
<td>56 (31-78)</td>
</tr>
<tr>
<td>JAMA</td>
<td>26</td>
<td>92 (24/26)</td>
<td>50 (29-71)</td>
</tr>
<tr>
<td>Obstetrics &amp; Gynecology</td>
<td>14</td>
<td>100 (14/14)</td>
<td>71 (42-92)</td>
</tr>
<tr>
<td>Ophthalmology</td>
<td>14</td>
<td>93 (13/14)</td>
<td>62 (32-86)</td>
</tr>
<tr>
<td>Overall</td>
<td>74</td>
<td>93 (69/74)</td>
<td>58 (46-70)</td>
</tr>
</tbody>
</table>

*CI indicates confidence interval.

*Odds ratios for the outcome variable (masking success) for each predictor variable, adjusted for all other variables.

Table 2.—Predictors of Masking Success: Multivariable Analyses of Reviewer Characteristics (at All Journals Except Journal of the American Geriatrics Society)*

<table>
<thead>
<tr>
<th>Type of Variable</th>
<th>Masked Reviewers (n = 136)</th>
<th>Unmasked Reviewers (n = 87)</th>
<th>Unadjusted OR (95% CI)†</th>
<th>Adjusted OR (95% CI)‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous Age, y</td>
<td>46</td>
<td>47</td>
<td>0.97 (0.94-1.00)</td>
<td>0.95 (0.89-1.02)</td>
</tr>
<tr>
<td>Reviewing experience, y</td>
<td>11</td>
<td>14</td>
<td>0.94 (0.91-0.98)</td>
<td>1.06 (0.97-1.16)</td>
</tr>
<tr>
<td>Articles reviewed in last 3 mo, No.</td>
<td>6.3</td>
<td>8.2</td>
<td>0.97 (0.93-1.00)</td>
<td>0.99 (0.95-1.05)</td>
</tr>
<tr>
<td>Articles published in last 5 y, No.</td>
<td>21</td>
<td>35</td>
<td>0.98 (0.97-0.99)</td>
<td>1.01 (0.99-1.03)</td>
</tr>
<tr>
<td>Time spent in research activities, %</td>
<td>32</td>
<td>48</td>
<td>0.98 (0.97-0.99)</td>
<td>1.01 (1.00-1.02)</td>
</tr>
<tr>
<td>Categorical Editorial experience (yes/no)</td>
<td>48</td>
<td>52</td>
<td>1.01 (0.62-1.65)</td>
<td>0.88 (0.43-1.75)</td>
</tr>
<tr>
<td>Sex (female/male)</td>
<td>16</td>
<td>21</td>
<td>1.47 (0.90-2.40)</td>
<td>0.73 (0.31-1.71)</td>
</tr>
<tr>
<td>Academic rank (full professor/other)</td>
<td>36</td>
<td>48</td>
<td>0.72 (0.43-1.20)</td>
<td>1.01 (0.41-2.49)</td>
</tr>
</tbody>
</table>

*Data expressed as mean number unless indicated otherwise. Odds ratios (ORs) for continuous variables represent the increase in odds associated with each unit increase in the continuous measure (eg, for age, the OR is for a 1-year change in age). The OR for an X% increase = (OR for unit increase)\(^{X}\).†Odds ratios for the outcome variable (masking success) for each predictor variable, unadjusted for any other variable. CI indicates confidence interval.

*Odds ratios for the outcome variable (masking success) for each predictor variable, adjusted for all other variables.

Reviewers were asked to provide information on age, sex, academic rank, reviewing experience, editorial experience, number of original research articles published in the last 5 years (including meta-analyses), number of manuscripts reviewed in the last 3 months, and percentage of time spent in research, clinical, and administrative activities.

Reviewers were asked to provide information on age, sex, academic rank, reviewing experience, editorial experience, number of original research articles published in the last 5 years (including meta-analyses), number of manuscripts reviewed in the last 3 months, and percentage of time spent in research, clinical, and administrative activities.

**Analysis**

Using the reviewer questionnaire, we collected data on masking success and reviewer familiarity with the author from all 7 journals and collected data on reviewer characteristics from 6 journals (not Journal of the American Geriatrics Society). (Copies of the questionnaire will be provided to interested readers on request.) We determined masking success by asking reviewers the question, “Do you think you can identify any of the principal authors of this manuscript?” If they answered yes, we instructed reviewers to list any authors they thought they could identify. If any reviewer correctly identified at least 1 author, we considered the reviewer unmasked to author identity. Thus, we measured masking success based on the percentage of reviewers who were able to guess author identity for reviewers who thought they knew at least 1 of the authors, we asked whether they were familiar with the work described in the manuscript before receiving it to review and whether the reviewer was familiar with the author’s work.

Reviewers were asked to provide information on age, sex, academic rank, reviewing experience, editorial experience, number of original research articles published in the last 5 years (including meta-analyses), number of manuscripts reviewed in the last 3 months, and percentage of time spent in research, clinical, and administrative activities.
research articles, and spent less time in research (Table 2, unadjusted ORs).

In a multivariable logistic regression model including reviewer age, sex, years of reviewing experience, number of articles published, number of articles reviewed, percentage of time spent in research, editorial experience, academic rank (full professor vs all other ranks), and journal (Annals of Emergency Medicine vs all other journals), we investigated the factors associated with correctly identifying an author. The unadjusted OR for masking success for reviewers at Annals of Emergency Medicine, compared with reviewers for other journals, was 4.8 (95% CI, 2.5-9.3; P<.001). However, after adjusting for the other 8 variables, the OR was 3.3 (95% CI, 1.4-7.2; P = .009), indicating a reduced, but still substantial, difference in masking success between Annals of Emergency Medicine and other journals. Thus, reviewer characteristics did not explain most of the difference in ability to identify authors between reviewers at Annals of Emergency Medicine and reviewers at other journals. Because factors such as age, number of publications, and reviewing experience are likely to be correlated, we performed regression analysis to examine which, if any, of the reviewer characteristics (other than the journal for which they were reviewing or the journal policy of masking) was an independent predictor of masking success. Of the reviewer characteristics we analyzed, percentage of time spent in research was the only significant independent predictor (P = .04; Table 2, adjusted ORs).

Because some manuscripts were reviewed by more than 1 reviewer, we repeated the logistic regression analyses, adjusted for clustering by manuscripts. The results were essentially the same as the unadjusted results (data not shown).

COMMENT

We confirmed the results of our previous study1 that the rate of masking success varies widely across journals, but found that a long-standing policy of masking did not increase masking success. Previous studies2-5 have found rates of masking success similar to those in this study. In our study, masking success was not associated with having a long-standing policy of masking, per se. In multivariable analysis, we found that masking success was associated with 1 reviewer characteristic. Reviewers spending a greater proportion of time in research were more likely to identify authors, suggesting that using reviewers with less research experience will increase masking success. Because research experience might be a desirable trait in reviewers for many manuscripts, however, excluding reviewers with this trait could be detrimental to the quality of reviewing. Thus, journals that mask author identity may have to balance their desire for improving the success of masking with the need for high-quality reviewers.

Characteristics of reviewers and other factors we examined did not completely explain differences in masking success. One possible factor might have been differences in masking procedures among journals. Although the minimum masking procedure followed a standard protocol, Annals of Internal Medicine and Annals of Emergency Medicine used additional procedures that may have affected masking success. Annals of Internal Medicine removed names and journal identification (but not titles or other reference information) from self-references in the text and reference section. However, the masking success rate at Annals of Internal Medicine was very close to the overall average. At Annals of Emergency Medicine, the instructions to authors stated that the journal’s policy was to mask author identity and requested that authors not include identifiers in running heads, but running heads would have been removed as part of normal study procedures. Although all 3 of the journals in our study with a policy of masking have had these policies for years, different journal “cultures” might have arisen that affected masking success (eg, authors at some journals might be less likely to self-reference).

Characteristics such as the age and size of medical specialties could influence the type of research performed and the probability that researchers will be able to identify others’ work. Contrary to what one might expect, however, JAMA and Annals of Internal Medicine, journals representing large and broad fields, had lower masking success rates than Annals of Emergency Medicine and Obstetrics & Gynecology. Our analyses suggest that to determine whether masking might be more successful in one field than another, specialty characteristics other than size of the field would need to be explored.

Our results suggest that reviewers’ research experience is associated with masking success, but that other unidentified factors are also involved. Indeed, results from a previous study1 suggested that author renown may in part affect masking success; this factor may be related to reviewer experience. In this study, we did not collect data on this author renown for all journals, and were thus unable to further investigate its effects. The study was also limited by the relatively small number of manuscripts enrolled at journals without a policy of masking. Additional research on a larger number of journals from a variety of medical specialties would be needed to determine what other characteristics of reviewers (eg, specialty or their relationships to authors), authors, journals, or medical specialties are associated with masking success.

Members of the PEER Investigators group include Michael Berkowitz, MD, University of Pennsylvania School of Medicine, Philadelphia; Phil B. Fontanarosa, MD, JAMA, Chicago, IL; Erica Frank, MD, MPH, Emory University School of Medicine, Atlanta, Ga; David Goldmann, MD, University of Pennsylvania School of Medicine, Philadelphia; Steven Goodman, MD, PhD, Johns Hopkins School of Medicine, Baltimore, Md; Roy Pitkin, MD, University of California, Los Angeles; and Rohit Varma, MD, University of Southern California School of Medicine, Los Angeles.

This work was supported in part by contract 467-MZ-501163 from the National Library of Medicine and the Agency for Health Care Policy and Research, and National Research Service Award 5 T32 HS00009 from the University of Pennsylvania.

We thank Marian Wiseman and Margaret Levene (Annals of Emergency Medicine), Mary Beth Schaeffer (Annals of Internal Medicine), Tracy Becker (Epidemiology), Elizabeth Webb (Journal of the American Geriatrics Association), Sharon Iversen (JAMA), Shani Koch (Obstetrics & Gynecology), and Donald Minckler, Sue Gertson, and Ann Dawson (Ophthalmology) for their help in managing the study and collecting data. We also thank Kenneth Rothman, PhD; Cristina Cumm, PhD (Epidemiology), and William Applegate, MD (Journal of the American Geriatrics Association) for their help in designing the study and instruments and collecting data.

References


JAMA. July 15, 1998—Vol 280, No. 3

Masking Author Identity in Peer Review—Cho et al

©1998 American Medical Association. All rights reserved.
In Reply.—Rexall Showcase International is a reputable direct sales company that sells unique products, including patented and proprietary food and nutritional supplement formulations, through a sales force of independent distributors, some of whom are physicians. Physicians who recommend RSI products have determined that these products provide optimum nutritional and preventive health benefits. Their decision to recommend RSI products, rather than directing their patients to purchase supplements and glean information from labels or nonphysicians, such as store clerks, is well-founded. Not all supplements “are created equal” and even well-known brands may be subpotent or may not dissolve properly.1 Rexall Showcase International fully supports 2 important corollaries to this axiom. First, a physician should only recommend a product if the physician believes, based on professional judgment, that the product will improve or maintain the patient’s health. Second, a patient’s decision about health and treatment alternatives, including the use of supplements, should be based on full disclosure. Therefore, we direct physicians to disclose to patients that they are RSI distributors and receive a monetary benefit from the sale of RSI products.

The letter from Dr Davis and colleagues is troubling because it appears that their institution has singled out as unethical only the direct sales of nutritional supplements (and particularly those of RSI). In many other contexts, physicians who recommend health supplies are entrusted to do so even though their compensation is affected. Many consumers in their capacity as patients have purchased medical devices or products from a physician and paid a premium price. More important, if physicians were barred from selling health-related items in their offices, patients would be prohibited from receiving vital information that physicians may believe, ethically, they have an obligation to disseminate.

Other factual inaccuracies or incomplete statements by Davis et al evidence a further failure to fairly evaluate RSI, its industry, and its products. For example, while the FDA did order certain guar gum pills off the market in 1990, it subsequently issued a rule that allows the over-the-counter marketing of products containing guar gum when they are taken with sufficient fluid.2 Rexall Showcase International has affirmed that the use of guar gum in its popular Bios Life 2 Fiber and Nutrient Drink Mix is generally recognized as safe, and the product is labeled with appropriate directions to ensure safe use. Contrary to other contentions, the claims made that RSI’s Bios Life 2 fiber matrix may lower blood cholesterol levels and reduce the risk of heart disease are amply supported by significant scientific agreement standards,3 clinical studies,4 and other scientific data.5 We believe RSI products are priced fairly, and because of patent and other proprietary protection, there are no comparable products for price comparison. All RSI products are of high quality, subject to rigid quality control and assurances, and manufactured in accordance with applicable current good manufacturing practices.

The RSI’s nutritional products are lawful, safe products with adequate substantiation for their claimed usage. If a physician has determined that the use of an RSI product is in the best interests of the patient and makes appropriate disclosure to the patient, consistent with professional ethics and good patient care, we believe a physician should not be precluded from earning income simply because the product is sold via direct sales.

David J. Schofield
Rexall Showcase International
Boca Raton, Fla

CORRECTIONS
Acknowledgment Omitted.—In the article entitled “Does Masking Author Identity Improve Peer Review Quality? A Randomized Controlled Trial” published in the July 15, 1998, Peer Review theme issue of THE JOURNAL (1998;280:240-242), Michael L. Callaham, MD, University of California, San Francisco, School of Medicine, was inadvertently omitted from the list of PEER Investigators in the acknowledgment.