

**Dale J. Benos, Edlira Bashari, Jose M. Chaves, Amit Gaggar, Niren Kapoor, Martin LaFrance, Robert Mans, David Mayhew, Sara McGowan, Abigail Polter, Yawar Qadri, Shanta Sarfare, Kevin Schultz, Ryan Splittgerber, Jason Stephenson, Cristy Tower, R. Grace Walton and Alexander Zotov**  
*Advan Physiol Educ* 31:145-152, 2007. doi:10.1152/advan.00104.2006

**You might find this additional information useful...**

---

This article cites 50 articles, 33 of which you can access free at:

<http://ajpadvan.physiology.org/cgi/content/full/31/2/145#BIBL>

This article has been cited by 2 other HighWire hosted articles:

**Is Peer Review Censorship?**

A. Casadevall and F. C. Fang  
*Infect. Immun.*, April 1, 2009; 77 (4): 1273-1274.  
[\[Full Text\]](#) [\[PDF\]](#)

**Publication in the AJR: Critical Interactions among Authors, Reviewers, and Section Editors**

T. H. Berquist  
*Am. J. Roentgenol.*, November 1, 2008; 191 (5): 1291-1292.  
[\[Full Text\]](#) [\[PDF\]](#)

Medline items on this article's topics can be found at <http://highwire.stanford.edu/lists/artbytopic.dtl> on the following topics:

Physiology .. Exertion  
Medicine .. Peer Review  
Education .. Accreditation  
Medicine .. Exercise

Updated information and services including high-resolution figures, can be found at:

<http://ajpadvan.physiology.org/cgi/content/full/31/2/145>

Additional material and information about *Advances in Physiology Education* can be found at:

<http://www.the-aps.org/publications/advan>

---

This information is current as of July 4, 2009 .

## The ups and downs of peer review

**Dale J. Benos, Edlira Bashari, Jose M. Chaves, Amit Gaggar, Niren Kapoor, Martin LaFrance, Robert Mans, David Mayhew, Sara McGowan, Abigail Polter, Yawar Qadri, Shanta Sarfare, Kevin Schultz, Ryan Splittgerber, Jason Stephenson, Cristy Tower, R. Grace Walton, and Alexander Zotov**

*Departments of Physiology and Biophysics, Nutrition, and Vision Sciences, University of Alabama, Birmingham, Alabama*

Submitted 17 October 2006; accepted in final form 29 December 2006

**Benos DJ, Bashari E, Chaves JM, Gaggar A, Kapoor N, LaFrance M, Mans R, Mayhew D, McGowan S, Polter A, Qadri Y, Sarfare S, Schultz K, Splittgerber R, Stephenson J, Tower C, Walton RG, Zotov A.** The ups and downs of peer review. *Adv Physiol Educ* 31: 145–152, 2007; doi:10.1152/advan.00104.2006.— This article traces the history of peer review of scientific publications, plotting the development of the process from its inception to its present-day application. We discuss the merits of peer review and its weaknesses, both perceived and real, as well as the practicalities of several major proposed changes to the system. It is our hope that readers will gain a better appreciation of the complexities of the process and, when serving as reviewers themselves, will do so in a manner that will enhance the utility of the exercise. We also propose the development of an international on-line training program for accreditation of potential referees.

publication; blinding

SCIENTIFIC PEER REVIEW has been defined as the evaluation of research findings for competence, significance, and originality by qualified experts (11). These peers act as sentinels on the road of scientific discovery and publication. Their reviews attempt to ensure the quality of scientific information (49), an act essential to reducing misinformation and confusion. Although this process is now part and parcel of scientific publishing, it was not always thought to be necessary.

When the *Journal des Scavans*, the first collection of scientific essays, was introduced in 1665 by Denis de Sallo, there was no peer review process in place. De Sallo's goal was to simply report the findings of others rather than guarantee their results (43). It was not until 1731 that the Royal Society of Edinburgh published *Medical Essays and Observations*, the first peer-reviewed collection of medical articles (43). In the first volume, the editor distributed the submitted essays for review to individuals he considered to be "most versed in these matters (55)." The Royal Society of Edinburgh recognized that the stamp of peer review did not necessarily mean the work was better than non-peer-reviewed publications. The purpose of the journal was solely to disseminate creative and important ideas; a disclaimer was provided stating that peer review did not guarantee truthfulness or accuracy. As always, the submitting authors were ultimately responsible for the quality and veracity of their own research (43).

The present-day peer review system evolved from this 18th century process. However, there have been concerns raised about bias, fairness, unnecessary delay, and general ineffectiveness of the process. Despite these shortcomings, peer review provides a formal opportunity for authors to gauge reaction to their work as well as allowing for the possible detection and subsequent correction of errors or flaws in logic

prior to an article's appearance in the public domain. Thus, peer review infuses "added value" into a publication. To better understand peer review, we examined the current process of peer review as well as its pros and cons and discuss proposed alternatives or modifications. It is our hope that a thorough discussion of these issues will increase awareness of the nuances of peer review, thereby improving its utility and minimizing its weaknesses.

### *Development of Peer Review*

The development of peer review was gradual and somewhat haphazard (12). Different editors employed varying styles of peer review. For instance, the *Lancet*, pre-1976, did not implement peer review as they considered it unimportant. Some journals, such as the *Journal of the American Medical Association* (JAMA), sent their submissions through an internal review panel and, on rare occasions, would send manuscripts to outside experts (55). The *British Medical Journal*, however, sent every noneditorial submission to a recognized expert by at least 1893 (12). By the late 20th century, peer review became institutionalized and is currently utilized by most biomedical journals. During this time, many journals embraced peer review in response to the increased specialization within each area of research and the increased competition among journals for manuscript submission (55). Over time, most researchers came to view peer review as the imprimatur for research articles.

The modern peer review system has evolved from its 18th century roots, and there is considerable variation in its application (28, 40). In general, it now embodies a process of systematically distributing, evaluating, and reaching a consensus on the merits of submitted manuscripts as evidenced by publication acceptance or rejection. The cornerstones of this process are the editors and expert reviewers (35, 67).

### *Editor and Reviewer Responsibilities*

Editors and reviewers have ultimate authority over a manuscript's fate, with editors primarily directing manuscript management and reviewers conducting manuscript assessment (38). Editors direct the process by selecting reviewers and communicating with authors and reviewers as well as by making the final decision on publication (61). Reviewers are selected based on their expertise and availability. Editors seek to balance the needs of their readers to receive only the most relevant information while providing a level of manuscript evaluation to authors that promotes continued submission in the face of stringent acceptance criteria.

Expert reviewers focus on detecting technical and stylistic flaws within the manuscript, determining the novelty of the

study, and making a recommendation of acceptance, rejection, or revision. They examine technical attributes as well as scientific quality, clarity of presentation, and ethical validity. Ideally, they do so in a manner consistent with ethical practices and journal guidelines (7). Reviewers donate substantial amounts of time and energy, frequently reviewing for multiple publications without remuneration (48, 68). Reviewers must provide timely feedback to editors and, as consulted experts, are frequently the de facto determiner for manuscript publication acceptance or rejection (61).

The motivation for reviewers and editors to participate in the peer review process has also evolved with time (37, 43). In the beginning, reviewers and editors were motivated to contribute in return for prestige and fame. Today, this reward continues, augmented with the additional motive of determining the quality and direction of research in a particular area (37). With widespread acceptance of peer review in the scientific community, manuscripts are not held in high esteem if they do not first pass through this process.

#### *A Manuscript's Gauntlet*

There are two key acceptance/rejection strata a manuscript must navigate following submission: the editorial level and the reviewer level (28). At first submission, editors can summarily reject a manuscript as inappropriate for their target audience or for a variety of other reasons. Although statistics vary widely by journal, up to 10% of manuscripts are rejected at this point (28). If the manuscript is deemed suitable by the editor, it advances to the expert reviewer stage. The number of reviewers is typically limited, with journals averaging two reviewers per manuscript (67). The reviewer's comments and recommendations are returned to the editor, who makes the decision to accept or reject the manuscript, often relying solely on the reviewers' recommendations. The manuscript can be accepted without revision, accepted with revision, or rejected. If a revision is requested, there is no guarantee of acceptance, and the manuscript may be rejected again after revision. Final acceptance/rejection rates for publication vary widely between scientific journals; in some instances, manuscript rejection can be as high as 90% (28).

#### *Assessing the Benefits of Peer Review*

To assess the influence of peer review on manuscript quality, an attempt was made to quantitatively characterize changes to manuscripts submitted to the *Annals of Internal Medicine* by performing line-by-line comparisons of published articles with the submitted versions. Substantive changes to manuscripts occurred in five major categories: too much information, too little information, inaccurate information, misplaced information, and structural problems (54). Another study of articles in the same journal measured manuscript quality using a 34-question instrument (27). Forty-four blinded physicians and epidemiologists examined manuscripts using the questionnaire; 97% of questionnaire items indicated improvement after editing and peer review. Four of the thirty-four items showed statistically significant improvement: discussion of limitations, acknowledgement and justification of generalizations, strength or tone of conclusions, and the use of confidence intervals. Although reviewer perception of the overall quality of manuscripts did not change significantly, the objective measures

probed by this questionnaire indicated otherwise. While this study indicates that the process is effective, the reliability of the 34-question instrument was low, and the study failed to differentiate between the effects of peer review and the effects of copy editing (27). Beyond the reviewers and readers of a paper, Wellen (66) reported that the majority of authors felt that their manuscript's quality was enhanced after peer review. Specifically, authors felt that manuscript content improved 37.2% of the time, organization improved 22.6% of the time, and conclusions were clearer 18.2% of the time.

#### *Assessing the Imperfections of Peer Review*

Despite peer review's acceptance within the research community, concerns have been raised about its overall effectiveness. Criticisms directed at the peer review process include bias toward certain authors, inability to detect major flaws, unnecessary delays in publication, and inability to uncover corruption/scientific misconduct. These concerns, especially the last, have weakened the scientific community's faith in the review process (47).

The most troublesome flaws in the peer review process are exposed by the increasingly frequent cases highlighting the inability of peer review to expose and minimize legal and ethical problems. High-profile cases such as those discussed by Cantekin et al. (15) present a quagmire of issues including reviewer bias, reviewer conflict of interest, breach of confidentiality, disclosure of funding sources, intellectual property rights, and the proper venue for publication of dissenting viewpoints.

#### *Peer Review and Bias*

Bias may be defined as systematic prejudice that prevents the accurate and objective interpretation of scientific studies (57). Common perceived biases include those in status, gender, and research attitudes (47). For example, favoring prominent researchers from well-reputed institutions seemed to be confirmed by the much-cited study by Peters and Ceci (16), which blinded reviewers to published papers from well-known research groups and found high rates of manuscript rejection based on scientific grounds. However, a study conducted by the *Journal of Pediatrics* indicated that while there was an association between high institutional status and acceptance of brief reports, this relationship was not observed in the acceptance rates of regular articles (22).

The presence of gender bias among reviewers also has been debated. In 1990, Lloyd (46) conducted an experiment in which she requested reviews of identical manuscripts with authors' names that were obviously male or female. In this relatively small cohort, female-authored manuscripts were accepted significantly more often by female reviewers (62%) than by male reviewers (21%). Male reviewers did not discriminate between male- or female-authored manuscripts in terms of acceptance rate (20–30%). Female reviewers were also significantly less likely to accept the male version of the paper (10%) compared with the female version (62%). Gilbert et al. (23) assessed gender bias at *JAMA* in a retrospective cohort study of 1,851 articles. While gender differences in reviewer and editor practices were observed, there was no measurable effect of gender on the final recommendation for publication or on the ultimate acceptance of papers. Interest-

ingly, female editors had larger workloads and were more likely to reject manuscripts summarily. When examining sexism in the peer review process as applied to funding, Grant et al. (28) found that peer-reviewed Wellcome Trust and Medical Research Council grants were awarded to men and women at approximately the same rate.

The objectivity of editors and reviewers can also be jeopardized by ideological differences, avoidance of unconventional ideas, and conflicts of interest. Reviewers may allow their beliefs to influence their reviews (32). For example, two papers differing only in their conclusions about the effects of maternal availability during childhood were created (32). The likelihood of acceptance or rejection depended largely on the reviewers' views on the subject.

It has also been suggested that editors or reviewers look unfavorably upon manuscripts containing unconventional ideas. Two often-cited examples are Hans Krebs' description of the citric acid cycle and Barbara McClintock's description of mobile gene elements, both of which were rejected by *Nature* (39). Both authors later won the Nobel Prize for their respective findings. Avoidance of avant garde and controversial topics by reviewers and editors could hamper the advance of science. Unconventional ideas can prevail over time when based on good science, but the peer review system can, and perhaps should, make that process challenging.

Another source of bias in peer review is conflict of interest. By definition, a conflict of interest (COI) occurs when participants in the publication process have personal interests that could inappropriately influence their judgment, regardless of whether or not their judgment is actually affected (25). COIs arise from a variety of relationships including financial and personal considerations, intellectual passions, and academic competition. The International Committee of Medical Journal Editors has determined that of these, financial considerations are the most prevalent (21, 36). These instances may involve employment, consultations, stock ownership, honoraria, expert testimony, and funding (42). Authors, reviewers, and editors may be less critical of manuscripts describing a method or product in which they hold personal interest, and companies may attempt to block publications that reflect negatively on their products (20). The rise in academia-industry relationships contributes to an environment in which editors or reviewers may behave in ways that increase their personal or financial gains (10). Editors and reviewers may be less critical of research from friends, collaborators, or scientists that agree with their beliefs or research. The opposite can occur if a reviewer encounters a manuscript from an investigator who supports a competing idea or one that is personally disliked. In some cases, reviewers or editors might benefit by delaying publication of data from another laboratory if they work on closely related topics.

Because of the secretive nature of the current peer review system in which reviewers' identities are kept hidden from authors, COIs are difficult to detect. Any conflicts must be reported by reviewers themselves. In 2001, at the Fourth International Conference on Peer Review in Biomedical Publication, Fiona Godlee of Biomed Central stated that in some cases, open (i.e., unblinded) peer review has helped detect reviewers' undisclosed competing interests (51). The *British Medical Journal* is currently utilizing a system of signed reviews, and the only "adverse effects" reported involved

authors exposing reviewers' COIs and some reviewers opting not to participate (50).

#### *Fraud and the Peer Review Process*

A major perceived failure of the peer review system is its inability to detect fraud (1, 30, 31). Scientists and the general public expect a peer-reviewed publication to be free of dishonesty. However, peer review cannot necessarily ensure that a paper is truthful. It can only claim that it is worth publishing (19). The public disclosure of scientific transgressions by Korean stem cell biologist Wook Suk Hwang (33), who falsely claimed to have created 11 human embryonic stem cell lines, highlighted the failures of peer review in detecting intentional misconduct. In the case of Hwang, the fraud was wrought so skillfully that it is perhaps unfair to blame the reviewers for failure to detect it. Unfortunately, blatant fraud passes through as well. Jan Hendrick Schon was a prolific young researcher at Bell Laboratories, publishing 100 papers between 1998 and 2002 with virtually no assistance (45). Several of his papers had significant impact and were regarded as major breakthroughs in multiple areas of solid-state physics. Readers later found that reviewers had failed to uncover gross anomalies, such as noise-free electrical data and identical data representing separate experiments in the same paper. The fraud evaded both an internal peer review process at Bell Laboratories and the reviewers of high profile journals such as *Nature* and *Science*. An in-depth inquiry by Bell Laboratories later identified 16 cases of scientific misconduct among 24 allegations raised (6). While this inquiry was admirable, the fact remains that this fraud was not caught by the reviewers, resulting in a loss of credibility for the scientific community as a whole. Reviewers may benefit from a written cue to assess possibility of fraud (e.g., a question on the standard Review Form may read: "There is no indication of fraudulent data in this manuscript: yes, no"). Outright fraud is likely very rare and therefore may not warrant implementation of systems to safeguard against it.

The vast majority of researchers do not purposefully create fraudulent data for publication. Unfortunately, the identification by peer review of methodological and statistical shortcomings in manuscripts is also limited. For example, in a 1998 experiment, eight weaknesses were introduced intentionally into a research article (26). The manuscript was sent to over 200 reviewers who, on average, identified only 2 of these weakness (26). Callaham et al. (14) also reported that reviewers could not spot two-thirds of the major errors introduced into a fake manuscript. A "white paper" on the peer review process from the perspective of integrity has just been published (59).

Many factors seem to contribute to these inadequacies in the peer review process, not the least of which are time limitations and reviewer burnout. Reviewers sometimes base their judgments on cues that have only a weak relation to quality such as statistical significance, large sample size, complex procedures, so-called "negative" data, and obscure writing (3). Atkinson et al. (5) conducted an experiment to determine whether reviewers place too much emphasis on statistical significance. They prepared three versions of a bogus manuscript where identical findings differed only by the degree of statistical significance. The reviewers recommended rejection of the paper with non-significant findings three times as often as those with significant findings. Although Atkinson et al. (5) criticize reviewers

for being unduly influenced by minor or false cues, such as lack of statistical significance, they themselves do not acknowledge that nonsignificant results may negatively impact the novelty of an article, a critical assessment criterion for manuscript quality (7, 8). Regardless, reviewers are expected to identify flaws, suggest improvements, assess novelty, and improve the quality of the manuscript instead of letting false cues hinder their better judgment (7, 8). Obscure or complex writing on complicated procedures should not fool a reviewer into misjudging the scientific caliber of a manuscript (4). Unfortunately, these false cues and biases are common enough to warrant improvements in the conduct of peer review.

#### *Delay Caused by Peer Review*

Scientific research and peer review are costly and time consuming. A year or more may pass before enough data are generated to warrant submitting one's work for publication. A manuscript is rarely accepted for publication after the initial submission. Thus, authors invariably modify their manuscripts. After these modifications, the author(s) may resubmit the work and hope for a more favorable review. In rapidly moving fields, scientists may allocate funds to a particular line of experiments only to be scooped. This concern has been somewhat alleviated by the World Wide Web, as publication is expedited by its appearance online within hours or days of acceptance.

#### *The Need for Peer Review*

Despite the difficulties and flaws discussed above, the process of peer review serves an important role in scientific publication. Abolishment of the review process would have significant repercussions. One major consequence would be the lack of an opportunity to respond to criticisms raised by experts prior to publication. In our opinion, this alone is sufficient reason to preserve peer review. After initial publication, authors may not be motivated to clarify or substantiate certain aspects of their research. In one study, the primary authors of three manuscripts responded to <50% of comments and criticisms in an open-review format (34). In another study conducted by the *Medical Journal of Australia*, only 7 of 56 manuscripts were modified in response to comments made by internet readers of accepted (but not yet published) manuscripts (9). Moreover, only 2% of those reading posted manuscripts replied with specific comments. This low rate of reader feedback may lead to inadequate scrutiny before a study's findings become widely available to the scientific community. The loss of the "added value" provided by formal peer review could be monumental.

In no arena is the danger of abolishing peer review greater than in clinical care. With the increasing availability of primary research to the public, publication of articles that have not been properly scrutinized by experts could compromise the health and treatment of patients. Some groups even advocate a second line of peer review to transition more appropriately scientific findings into clinical practice guidelines (29). Loss of peer review in this environment would undercut important patient safety protections.

Finally, the loss of a formalized peer review process may possibly increase the number of fraudulent studies. While there is no significant research that speaks specifically to this issue, the loss of critical examination of a submitted manuscript by a

group of experts would remove a deterrent to the submission of dubious publications. In fact, Dr. Lawrence Altman, a science writer for the *New York Times*, states that, "editors of the journals and many scientists consider the system's expense and time consumption worthwhile in the belief that it weeds out shoddy work and methodological errors and blunts possible biases by scientific investigators" (2).

#### *Changes to Peer Review*

With peer review so ingrained in the publication process, it would be impractical, detrimental, and unwise to abolish it. The past few decades have seen many proposed changes, with blinding, unmasking, and open review being the most prominent.

Before embarking on a discussion of these changes, a common vocabulary should be established. Currently, in the most widely adopted form of peer review, reviewers of a manuscript are not blind to authors and their affiliations, although the reviewers' identities are masked from authors and from other reviewers. Blinding is then defined as the process of hiding the identity and affiliations of the authors of a manuscript. Unmasking is defined as identifying the reviewers to the authors or to the other reviewers. Reviewers can also sign their review, attaching their name or critique to a manuscript, making themselves known to the author. In the literature, blinding and masking are often used interchangeably. An "open" review process is one that allows a larger group of reviewers, generally described as the public, to comment upon and critique a manuscript that is already published, usually online. This can be coupled with unmasking to create a completely transparent process. We will discuss the purpose of these proposed modifications, perceived benefits, perceived difficulties, and relevant data and conclude with an analysis of possible modifications.

Blinding aims to remove reviewer bias from the peer review process. It is a logical extension of the blinding process used in experimental design. By removing the names and affiliations of the authors from a manuscript, any bias for or against an author or institution will be removed or reduced. Theoretically, Nobel laureates from elite institutions and budding researchers from lesser-known schools will be treated equally. Although this seems fair and logical, in practice it is very difficult to achieve complete blinding.

Data regarding blinding are somewhat conflicted. McNutt et al. (52) studied the effect of blinding reviewers to author identity and found that the quality of reviews, as measured by the editors, was improved by blinding reviewers. They also found that they were successful in blinding the reviewers 73% of the time. However, the improved quality of the reviews may have been due to the reviewers working more diligently because they were aware that they were involved in a research study. When van Rooyen et al. (64, 65) studied the effects of blinding and unmasking in a relatively large, randomly controlled trial, they found no significant differences in quality of reviews, recommendations for publication, or time taken to review the manuscript. However, in their cohort, removing the authors' details from the title page and acknowledgments led to successful blinding only 51.2% of the time (65). Self-referential writing and small research fields were the most commonly given reasons for unsuccessful blinding. A more rigorous study by Cho et al. (18) examining the success of blinding found that

removing only the authors' names and institutional affiliations was successful in blinding reviewers about 60% of the time. Interestingly, compared with the reviewers who were able to identify the authors despite blinding, those reviewers who were successfully blinded to authors spent less time conducting research, published less often, and had less experience reviewing, qualities that suggest inexperience or unfamiliarity with the research area. In 1984, the journal *Medicine and Science in Sports and Exercise* evaluated the responses of 222 authors (111 accepted and 111 rejected manuscripts) on the blind review system (62). They found that the majority of authors favored the blind review system. Notably, the authors indicated that their present and past status as an investigator at their current or previous location had little influence on the editorial decision.

Sidestepping the issue of successful blinding, Godlee et al. (26) modified a previously published article by intentionally introducing weaknesses and examined the effects of blinding reviewers or authors and the signing of reviews on the ability of a referee to detect errors. They found no significant difference between the qualities of the reviews, as measured by the number of errors found, but found that blinded reviewers were more likely to recommend the weakened article be published. However, the reviewers were not experts on the subject, and this fact could have confounded the results.

Although blinding manuscripts is difficult, removing identifying material from an abstract is relatively simpler as there are fewer concerns regarding self-referential citations. Ross et al. (56) studied the effect of blinding on abstract acceptance to the American Heart Association's annual research meeting. They found that when reviewers were not blinded to authors' identity, there was a bias for authors from American institutions, those from prestigious institutions, and authors from countries where English was the official language. By blinding reviewers, these biases were significantly attenuated. However, the review process for abstracts and manuscripts is very different. Much like Ross' work, a study conducted at the *Journal of Pediatrics* indicated an association between high institutional status and acceptance of brief reports. However, this relationship was not observed for acceptance rates of major papers (22). Regardless, these results do suggest that blinding may be useful in reducing bias, at least in some circumstances.

Perhaps the truest test of the quality of an article lies in how often it is cited, a reasonable proxy for the impact of the paper on the scientific community. A retrospective study (44) of 28 economic journals from 1984 to 1989 found that articles published in a blinded review process were cited more than articles published in a nonblinded review process. However, interpreting this study is difficult because blinding success rates were unknown.

Although data regarding the effects of blinding are conflicted, the logic behind blinding appears solid. While scientists strive to be objective, they are human. Removing identifiers such as names and institutions should minimize the bias of interpersonal relationships. The major negative aspect of blinding, as shown by Godlee et al. (26), is that reviewers may be more likely to recommend publication of a poor quality manuscript when they are blinded. To make the process of blinding more successful, appropriate procedural changes can be made. Authors should be directed to minimize self-referential writing, as it appears to be one of the most common reasons for

unsuccessful blinding (65). It should be relatively straightforward to design software that can parse a text, highlight phrases that are self-referential (such as "we have shown"), and alert an editor to a possible cause of unblinding. Some journals have already assimilated blinding into their regular practices. For example, the *American Journal of Roentgenology* uses author and referee blinding in reviewing its >1,700 yearly submitted manuscripts (<http://www.ajronline.org/>).

It should also be noted that blinding does not alleviate all types of publication bias. Removing an author's name cannot remove biases against unconventional methodology, radical new ideas, negative results, or results that contradict a reviewer's viewpoints. While blinding articles is a simple step that could be taken to reduce bias and slightly improve review quality, the literature does not suggest that it is required for quality review of manuscripts (22). Furthermore, more data are needed to determine whether or not blinding could have the unintended consequence of hiding conflicts of interest from reviewers. Should a possible COI on the part of the authors be revealed to reviewers and if so when? Research by Chaudhry et al. (17) shows that a statement of competing interests leads readers to feel that an article is less interesting, important, relevant, valid, and believable than that same article without a COI statement. As a growing number of researchers and institutions begin patenting intellectual property for future commercial applications, it is imperative that the scientific community address how COI statements should be handled in the peer review process. The federation of American Societies for Experimental Biology has made a major step in this regard by releasing a white paper on COI ([http://opa.faseb.org/pdf/FASEB\\_COI\\_paper.pdf](http://opa.faseb.org/pdf/FASEB_COI_paper.pdf)).

Unmasking identifies reviewers to the authors or to the other reviewers. On the positive side, this would reduce the ability of reviewers to hide COIs and would increase the transparency of the entire process because reviewer identity would at least be known to another reviewer. If reviews were signed, the reviewers would open themselves to criticism and likely conduct a more thorough review. In some cases, reviewer COIs could be detected more easily. On the other hand, unmasking or signing of reviews may also cause reviewers to be less critical of some works. Older, more established authors could enact retribution on younger, less established reviewers who were perceived as being overly critical. Thus, masking reviewers may serve to protect the reviewers and the process of review.

As mentioned earlier, van Rooyen et al. (64, 65) studied the effects of unmasking the identity of a reviewer to their co-reviewer. Their group found no significant differences in the quality of reviews as assessed by editors. In another study, the effects of unmasking reviewers to the authors of a manuscript was examined (63). Again, there were no significant differences in the quality of reviews. However, reviewers were 12% more likely to decline to review if the process were unmasked. When surveying authors of manuscripts, 55% were in favor of asking reviewers to sign, whereas only 26% were against it (with the rest being indifferent). When Godlee et al. (26) studied the effect of having reviewers sign their names to reviews, there were no significant differences in the quality of reviews as assessed by their ability to detect errors in the manuscript. However, in this study and in another unpublished pilot study, half of the reviewers declined to sign their name when asked to. Another study showed that reviewers who

signed their reviews were more courteous and constructive in their criticism, as assessed by editors (52). Manuscript authors observed no significant differences except that they perceived the signed reviews to be slightly fairer. However, signers were more likely to recommend publication and rate the manuscript higher than nonsigners (52). A survey of reviewers for *Medical Education* found that while 73% were not averse to signing their reviews, only a minority actually signed (60). First-time reviewers were also more likely to be unwilling or unsure of signing, possibly suggesting a lack of confidence in the quality of their reviews or fear of possible retribution due to a negative critique.

The data suggest that unmasking reviewers or having them sign their reviews has no beneficial effects in terms of quality or strength of the review. It does, however, make reviewers more likely to decline to review a manuscript. Although scientists may not have a problem with signing reviews in surveys, reviewers appear to prefer signing a review recommending publication (52, 60). This adds credence to the idea that signed reviews correlate with positive reviews. With that stated, more transparency in the review process may make authors feel better, but the lack of any measurable quality benefit and the increased difficulty in securing expert reviewers outweigh any perceived benefit.

An open review process serves two main functions: it makes research available immediately and it allows multiple people to comment upon a manuscript. Proponents of this method state that open review will increase the quality of manuscripts, disseminate information faster, and reduce the burden on reviewers. Currently, some journals have adopted or are experimenting with this method of review. For example, *Atmospheric Chemistry and Physics* (ACP) uses an open two-stage review process in which a manuscript is made available online immediately, commented upon by designated reviewers and interested scientists, revised if needed, and then sent through a shorter, traditional review process (41). This open review process, while transparent, still allows reviewers to choose to be anonymous, with about two-thirds of the reviewers remaining hidden. However, scientists choosing to comment on the paper during the first stage must sign their comments. The editors of this journal report that, despite their low rejection rate of ~20%, they are ranked 12th of 169 journals in the combined areas of meteorology and atmospheric sciences and environmental sciences according to impact factor (41). This appears to support the argument that either higher quality manuscripts are submitted to a journal utilizing an open review process or that the process itself improves the manuscript significantly. However, the appropriate experimental design has not been applied to test rigorously this supposition.

The journal *Electronic Transactions on Artificial Intelligence* (ETAI) also divides the peer review process into two stages: a completely transparent reviewing stage and a refereeing stage (58). As soon as it is decided that the paper falls within the scope of ETAI's readership, the manuscript is published electronically by an organization affiliated with ETAI. This protects the authors from theft of their intellectual property. There is at least a 3-mo period during which members of the community can comment on the paper. Unlike ACP, all commenters/reviewers must sign their comments. These comments are then screened by an editor to ensure that they are relevant and pass a basic quality threshold. After this discus-

sion period, authors may revise the paper, and it can then be submitted to the referees, who decide whether or not the paper should be published with a simple pass/fail. If it passes, the article is published in an issue of the journal. If it fails, it remains published electronically, but the article is not published by the journal. According to the general editor of ETAI, Erik Sandewall, authors report favorably on the process (58). Criticism, positive or negative, is welcome because it draws attention to the manuscript. It is problematic, however, that two versions of basically the same work exist in the published record.

Both of these journals (ACP and ETAI) do not unmask the people who decide whether or not a paper is publication worthy. They allow for signed comments and signed peer review, but they do not ask referees to cast open votes upon articles. This does not remove any bias, perceived or real, by referees or editors. Thus, these forms of open review, while alleviating delays and increasing transparency, will not attenuate perceptions of bias at the actual acceptance step of the process.

#### *The Future of Peer Review*

Well-organized forms of free information exchange exist today and may provide useful models for scientific publication. Wikipedia, an online encyclopedia that allows individuals to submit and edit information on any topic of their choosing, offers an interesting case study on this topic. Because entries in Wikipedia can be altered by anyone with access to the internet, one would reason that the information contained within it may lack a significant degree of accuracy. However, it was found to have roughly the same percent accuracy as Encyclopedia Britannica, which employs a board of 14 editors, all of whom are considered experts in their respective fields (21, 24).

Biology Direct is another example of an alternative peer review system. The system is faster than the conventional system: a paper is published once three editors agree to formally critique it. Its transparency reveals COIs, because reviewers' names and comments are published alongside the article. This transparency allows readers to judge for themselves whether poor reviews were deserved. It also enables authors to defend their work, because they can post rebuttals alongside the reviews. Additionally, Biology Direct encourages controversial publications. Therefore, Biology Direct is an example of an alternative to peer review that is currently being used to publish original biomedical research and may provide some clues about how to improve the conventional system of peer review (<http://www.biology-direct.com>).

Ideally, a process that minimizes bias, promotes discussion, reduces time to publication, decreases variability in the peer review process, and increases overall quality of work without stifling new and radical ideas should be adopted. At the very least, training of reviewers in the nuances of how to review a manuscript and provide useful critiques to both the authors and editors should be formally instituted for every journal. Designing a structured, standardized training course for reviewers would define protocols for reviewing, proper criteria to apply, and common pitfalls to avoid. Several journals, such as *Annals of Emergency Medicine*, the *American Journal of Roentgenology*, the *American Journals of Physiology*, and the *British Medical Journal* ([http://www.bmj.com/advice/peer\\_review/](http://www.bmj.com/advice/peer_review/)) have available training material for their reviewers (8, 13, 53).

However, a simple training protocol will likely not be enough to ensure that reviews remain standardized and unbiased. Mock papers at regular intervals can serve to easily assay the quality of reviews, and regular quizzes or newsletters can be used to reinforce important key points. This relatively simple process would likely resolve many difficulties that are encountered by editors, reviewers, and authors. Nonetheless, we propose that organizations of biomedical journal editors, like the International Committee of Medical Journal Editors or the Council of Science Editors, develop an on-line reviewer accreditation course, following the template already prepared by the *British Medical Journal*.

An attempt to design an ideal peer review process must accept a priori that manuscripts will vary in quality and impact. However, instead of attempting to bury the lower end of the curve by rejecting manuscripts, the goal should be to shift the distribution toward higher quality manuscripts. To this end, a hybrid open review process may be a solution. Erik Sandewall, editor of *ETAI*, notes that the issue of intellectual property theft must be addressed in a true open-review process (58). However, most journals frown on republishing data that are already publicly available. This problem can be sidestepped by limiting access to a manuscript to the entire cohort of reviewers and editors of a journal. The entire group could comment on any article that piques their interest while a group of three to five reviewers assigned by an editor would provide a formal review. Those choosing to comment should be allowed anonymity if they desire while the formal reviewers must always be allowed anonymity. However, both comments and formal reviews should be available to all members of the reviewing group. The formal review process should provide feedback in written format while also providing numeric scores to track reviewer scoring patterns. Authors would be able to access and respond to comments and reviews of their manuscripts, with the handling editor moderating and ultimately adjudicating the decision.

This system will allow for helpful discourse in a relatively open format while providing referees anonymity to critique papers fully and without repercussion. Authors will be able to discuss their papers in an open forum and directly interact with reviewers. If so desired, editors may choose to strip the article of identifying content, such as names, institutions, and self-referential statements, to attempt to blind the manuscript to the reviewing board. If blinding is desired, authors should be instructed to minimize self-referential text. During the review process, the authors could be assigned aliases and have their comments reviewed to ensure that their identities are not revealed during the discussion. While some manuscripts may not garner as many comments as others, editors may be able to assess manuscript importance based on interest raised during the publishing process as well as comments of reviewers. Also, using a numeric scoring system will allow editors to decrease variability and clarify scoring criteria when a reviewer's scores are consistent outliers. It will allow editors to evaluate editorial board member participation, timeliness, and quality. These simple changes would make the process more transparent to authors and hopefully increase the overall quality of published manuscripts. American Journal of Physiology publications would be an ideal place to experiment with such a plan, because one or more of its journals can be used to field test a new review process.

## Conclusions

The process of peer review is steeped in tradition. In the realm of innovation and science, tradition operates on a day-to-day basis. The current system of peer review is not perfect, and, while most scientists believe it is necessary, indeed, desirable, the core assumptions inherent in the process must be evaluated and adapted to the changing environment. While research has examined modifications to the process, biases, and other flaws, perhaps the purpose of peer review and some key assumptions should be examined. For example, should manuscript peer review emphasize the validity of the science or its merit? While peer review is often viewed as the gatekeeper to the realm of truth, a staff of editors and reviewers cannot make that distinction in a few months. In terms of interesting assumptions overlooked, does the process of peer review require a true peer or expert in the field? Would unfettered publishing of findings lead to more efficient or faster progress in science? What does the stamp of peer review mean to the readers of the literature? We hope that this essay serves to incite discussion upon the peer review debate.

## ACKNOWLEDGMENTS

We thank Margaret Reich for invaluable comments on multiple drafts of this article and Rebecca Todd for editorial assistance.

## GRANTS

This work was supported in part by National Institutes of Health Grants DK-037206, CA-101952, and DK-53090.

## REFERENCES

1. **Allchin D.** Reassessing van Helmont D, reassessing history. *Bioscience. J Coll Biol Teaching* 19: 3–5, 1993.
2. **Altman L.** When peer review produces unsound science. *The New York Times*, 2002, p. D6 and F6.
3. **Armstrong J.** Barriers to scientific contributions: the author's formula. *Behav Brain Sci*: 197–199, 1982.
4. **Armstrong JS.** Peer review for journals: evidence of quality control, fairness, and innovation. *Sci Eng Ethics* 3: 63–84, 1997.
5. **Atkinson DR, Furlong BE, Wampold MJ.** Statistical significance, reviewer evaluations, and the scientific process: is there a statistically significant relationship? *J Counsel Psychol* 29: 189–194, 1982.
6. **Bell Labs.** *Report of the Investigation Committee on the Possibility of Scientific Misconduct in the Work of Hendrick Schon and Coauthors.* Madison, WI: Bell Labs, 2002.
7. **Benos DJ, Fabres J, Farmer J, Gutierrez JP, Hennessy K, Kosek D, Lee JH, Olteanu D, Russell T, Shaikh F, Wang K.** Ethics and scientific publication. *Adv Physiol Educ* 29: 59–74, 2005.
8. **Benos DJ, Kirk KL, Hall JE.** How to review a paper. *Adv Physiol Educ* 27: 47–52, 2003.
9. **Bingham C, van der Weyden MB.** Peer review on the Internet: launching eMJA peer review study 2. *Med J Aust* 169: 240–241, 1998.
10. **Bradley SG.** Managing competing interests. In: *Scientific Integrity: Text and Cases in Responsible Conduct of Research* (3rd ed.), edited by Macrina FL. Washington, DC: ASM, 2005.
11. **Brown T.** *Sense About Science. Peer Review and the Acceptance of New Scientific Ideas* (online). <http://www.senseaboutscience.org.uk/pdf/PeerReview.pdf> [30 July 2006].
12. **Burnham JC.** The evolution of editorial peer review. *JAMA* 263: 1323–1329, 1990.
13. **Callaham ML, Schriger D, Cooper R.** *An Instructional Guide for Peer Reviewers of Biomedical Manuscripts* (online). <http://www3.us.elsevierhealth.com/extractor/graphics/em-acep/index.html> [5 December 2006].
14. **Callaham ML, Baxt WG, Waeckerle JF, Wears RL.** Reliability of editors' subjective quality ratings of peer reviews of manuscripts. *JAMA* 280: 229–231, 1998.

15. **Cantekin EI, McGuire TW, Potter RL.** Biomedical information, peer review, and conflict of interest as they influence public health. *JAMA* 263: 1427–1430, 1990.
16. **Ceci SJ, Peters DP.** Peer review—a study of reliability. *Change* 14: 44–48, 1982.
17. **Chaudhry S, Schroter S, Smith R, Morris J.** Does declaration of competing interests affect readers' perceptions? A randomised trial. *Br Med J* 325: 1391–1392, 2002.
18. **Cho MK, Justice AC, Winker MA, Berlin JA, Waeckerle JF, Callahan ML, Rennie D.** Masking author identity in peer review: what factors influence masking success? PEER Investigators. *JAMA* 280: 243–245, 1998.
19. **Couzin J.** Stem cells and how the problems eluded peer reviewers and editors. *Science* 311: 23–24, 2006.
20. **Deyo RA, Psaty BM, Simon G, Wagner EH, Omenn GS.** The messenger under attack—intimidation of researchers by special-interest groups. *N Engl J Med* 336: 1176–1180, 1997.
21. **Encyclopaedia Britannica.** *The Encyclopaedia Britannica Editorial Board of Advisors* (online). <http://corporate.britannica.com/board/> [28 February 2007].
22. **Garfunkel JM, Ulshen MH, Hamrick HJ, Lawson EE.** Effect of institutional prestige on reviewers' recommendations and editorial decisions. *JAMA* 272: 137–138, 1994.
23. **Gilbert JR, Williams ES, Lundberg GD.** Is there gender bias in JAMA's peer review process? *JAMA* 272: 139–142, 1994.
24. **Giles J.** Internet encyclopaedias go head to head. *Nature* 438: 900–901, 2005.
25. **Godlee F.** The ethics of peer review. In: *Ethical Issues in Biomedical Publication*, edited by Jones AH, McLellan F. Baltimore, MD: Johns Hopkins Univ. Press, 2000.
26. **Godlee F, Gale CR, Martyn CN.** Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. *JAMA* 280: 237–240, 1998.
27. **Goodman SN, Berlin J, Fletcher SW, Fletcher RH.** Manuscript quality before and after peer review and editing at *Annals of Internal Medicine*. *Ann Intern Med* 121: 11–21, 1994.
28. **Hargens LL.** Variation in journal peer review systems. Possible causes and consequences. *JAMA* 263: 1348–1352, 1990.
29. **Haynes RB, Cotoi C, Holland J, Walters L, Wilczynski NL, Jedraszewski D, McKinlay J, Parrish R, McKibbin A for the McMaster Premium Literature Service Project.** A second order of peer review: a system to provide peer review of the medical literature for clinical practitioners. *JAMA* 295: 1801–1808, 2006.
30. **Hershey D.** Misconceptions about Helms' willow experiment. *Plant Sci Bull*: 78–84, 2003.
31. **Hershey D.** The widespread misconception that the tambalacoque or calvaria tree absolutely required the dodo bird for its seed to germinate. *Plant Sci Bull*: 105–108, 2004.
32. **Hojat M, Gonnella JS, Caelleigh AS.** Impartial judgment by the "gatekeepers" of science: fallibility and accountability in the peer review process. *Adv Health Sci Educ Theory Pract* 8: 75–96, 2003.
33. **Holden C.** Stem cell research: Korean cloner admits lying about oocyte donations. *Science* 310: 1402–1403, 2005.
34. **Horton R.** Postpublication criticism and the shaping of clinical knowledge. *JAMA* 287: 2843–2847, 2002.
35. **Ingelfinger FJ.** Peer review in biomedical publication. *Am J Med* 56: 686–692, 1974.
36. **International Committee of Medical Journal Editors.** *Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication* (online). <http://www.icmje.org/index.html#ethic> [30 July 2006].
37. **Isohanni M.** Peer review—still the well-functioning quality control and enhancer in scientific research. *Acta Psychiatr Scand* 112: 165–166, 2005.
38. **Kassirer JP, Campion EW.** Peer review. Crude and understudied, but indispensable. *JAMA* 272: 96–97, 1994.
39. **Kilwein JH.** Biases in medical literature. *J Clin Pharm Ther* 24: 393–396, 1999.
40. **Knoll E.** The communities of scientists and journal peer review. *JAMA* 263: 1330–1332, 1990.
41. **Koop T, Pöschl U.** An open, two-state peer-review journal. *Nature* (online). <http://www.nature.com/nature/peerreview/debate/nature04988.html> [30 July 2006].
42. **Krimsky S, Rothenberg LS.** Financial interest and its disclosure in scientific publications. *JAMA* 280: 225–226, 1998.
43. **Kronick DA.** Peer review in 18th-century scientific journalism. *JAMA* 263: 1321–1322, 1990.
44. **Laband DN, Piette MJ.** A citation analysis of the impact of blinded peer review. *JAMA* 272: 147–149, 1994.
45. **Lerner EJ.** Fraud shows peer review flaws. *Industrial Physicist*: 12–17, 2003.
46. **Lloyd ME.** Gender factors in reviewer recommendations for manuscript publication. *J Appl Behav Anal* 23: 539–543, 1990.
47. **Lock S.** Does editorial peer review work? *Ann Intern Med* 121: 60–61, 1994.
48. **Lock S, Smith J.** What do peer reviewers do? *JAMA* 263: 1341–1343, 1990.
49. **Macrina FL.** Authorship and peer review. In: *Scientific Integrity: Text and Cases in Responsible Conduct of Research* (3rd ed.), edited by Macrina FL. Washington, DC: ASM, 2005.
50. **McCook A.** Is peer review broken? *The Scientist* 20: 26, 2006.
51. **McLellan F.** Peer-review meeting participants urge greater accountability. *Lancet* 358: 991, 2001.
52. **McNutt RA, Evans AT, Fletcher RH, Fletcher SW.** The effects of blinding on the quality of peer review. A randomized trial. *JAMA* 263: 1371–1376, 1990.
53. **Provenzale JM, Stanley RJ.** A systematic guide to reviewing a manuscript. *J Nucl Med Technol* 34: 92–99, 2006.
54. **Purcell GP, Donovan SL, Davidoff F.** Changes to manuscripts during the editorial process: characterizing the evolution of a clinical paper. *JAMA* 280: 227–228, 1998.
55. **Rennie D.** Editorial peer review: its development and rationale. In: *Peer Review in Health Sciences*, edited by Godlee F, Jefferson T. London: BMJ Books, 1999.
56. **Ross JS, Gross CP, Desai MM, Hong Y, Grant AO, Daniels SR, Hachinski VC, Gibbons RJ, Gardner TJ, Krumholz HM.** Effect of blinded peer review on abstract acceptance. *JAMA* 295: 1675–1680, 2006.
57. **Rumsey TS.** One editor's views on conflict of interest. *J Anim Sci* 77: 2379–2383, 1999.
58. **Sandewall E.** Systems: opening up the process. *Nature*, 2006.
59. **Scott-Lichter D.** *aEPC, Council of Science Editors. CSE's White Paper on Promoting Integrity in Scientific Journal Publications* (online). <http://www.nature.com/nature/peerreview/debate/nature04994.html> [30 July 2006].
60. **Snell L, Spencer J.** Reviewers' perceptions of the peer review process for a medical education journal. *Med Educ*: 90–97, 2005.
61. **Stehbens WE.** Basic philosophy and concepts underlying scientific peer review. *Med Hypotheses* 52: 31–36, 1999.
62. **Tipton CM.** *Medicine and Science in Sports and Exercise—1984. Med Sci Sports Exerc* 16: viii, 1984.
63. **van Rooyen S, Godlee F, Evans S, Black N, Smith R.** Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial. *Br Med J* 318: 23–27, 1999.
64. **van Rooyen S, Godlee F, Evans S, Smith R, Black N.** Effect of blinding and unmasking on the quality of peer review. *J Gen Intern Med* 14: 622–624, 1999.
65. **van Rooyen S, Godlee F, Evans S, Smith R, Black N.** Effect of blinding and unmasking on the quality of peer review: a randomized trial. *JAMA* 280: 234–237, 1998.
66. **Weller AC.** A comparison of authors publishing in two groups of U. S. medical journals. *Bull Med Libr Assoc* 84: 359–366, 1996.
67. **Weller AC.** Editorial peer review in US medical journals. *JAMA* 263: 1344–1347, 1990.
68. **Yankauer A.** Who are the peer reviewers and how much do they review? *JAMA* 263: 1338–1340, 1990.