The Impact of Peer Review on Intellectual Freedom

MARY BIGGS

ABSTRACT
The nature and history of peer review are described and its positive and negative effects considered. It is concluded that although peer review tends to penalize innovation and nonconformity, it is indispensable to scholarly publishing.

INTRODUCTION
"Peer review" connotes genteel collegial cooperation, while "refereeing" suggests the boxing ring, the football field, the objective mediator pressured by impassioned opponents. Yet the terms are often used synonymously as they will be in this discussion. Though both collegial and objective in theory, in practice the process is corruptible by ignorance, timidity, envy, greed, bias, and other common sins. It is this gap between the ideal and the real, coupled with peer review's extraordinary impact on scholars' professional futures and immediate feelings, that makes it so controversial. Yet the real danger, and strength, of peer review lies not in its consequences for the authors reviewed but for prospective readers of their work. Can peer review, which should help protect access to sound ideas, actually impede access? If so, under what conditions, and how can these be prevented?

Before tackling these questions, we must understand the nature of the process.

HISTORY AND NATURE OF PEER REVIEW
Essentially peer review means what it says: the review of a person's work by one or more people qualified to be called professional peers. When restricted to the evaluation of research and writing, peer review
involves the scrutiny of grant proposals, or of article or book manuscripts, by two or more people with suitable subject and methodological expertise. These people, individually, then recommend acceptance, revision, or rejection to whomever controls the process—usually a grants administrator or editor. The recommendations are handled in various ways: from automatic adoption of the majority's view to careful study of each recommendation as advisory only, followed by a relatively independent final judgment. It all depends on the particular journal's or publisher's or agency's selection policy, which isn't necessarily clearly communicated to the public.

Here the focus will be on peer review in journal publishing, which differs in its details, though perhaps not in the broad issues raised, from peer review in grant funding and book publishing.

Several scholars have described the history of scholarly journals, including, with admirable concision, sociologists Harriet Zuckerman and Robert K. Merton (1971). According to them, seeds of the peer review process were sown with the founding in 1665 of the first English-language scientific journal, *Philosophical Transactions*. The Council of Britain's Royal Society, which instituted the journal, stipulated that prior to publication, each monthly issue should be "reviewed by some of the [Council] members." The purpose, then as now, was to guard quality (p. 69). Over the centuries, journal peer review was systematized but never made uniform and was widespread but never universal. It is used in all disciplines, though with considerable variation in underlying assumptions, implementation, and results. And, despite the long history supporting peer review, its value continues to be debated.

Several journals, most of them in science and medicine, have published detailed explanations of their peer review practices (e.g., see Editorial staff, 1988, pp. 412-14; Enos, 1987; Carney & Lundberg, 1987, p. 87; Lundberg & Carney, 1986, p. 3286; Stossel, 1985, pp. 658-59; Rubin & Carroll, 1981, pp. 103-04; "The Refereeing System...", 1978, pp. 9-10). In a succinct monograph, Stephen Lock (1985) synthesized every substantial publication through the early 1980s dealing with peer review in medical journals. And in 1978, Michael Gordon (1978) reported his thorough survey, based on interviews with editors, of peer review methods used by thirty-two London-based research journals in several disciplines. Title by title, he set forth his findings. The validity of many of these data depends, of course, on the honesty, clear-sightedness, and comprehensiveness of the editors' presentations. Still, a good deal of anecdotal information and some solid research findings have appeared in print.

To summarize: manuscripts are checked in; are usually screened—cursorily or carefully—in-house; and some percentage—ranging from
less than half to nearly all—is sent on to two or more referees. They are selected according to information gathered through every conceivable means, from personal knowledge to "invisible college"-generated referrals to literature reviews to broad-scale questionnaires that enables "profiling" and the building of a formal referee database. And the information is stored in every possible way, from the editor's memory to a card file to a computer. Accompanying each review request may be detailed evaluation guidelines, a referee's report form to fill out, both, or neither. Some sort of review deadline is likely to be specified, though it may not be enforced. The time allowed to the referee varies. Authors' names may or may not be disclosed to referees. Referees' comments may be passed on to authors whole and unrevised, excerpted, paraphrased, or not at all. When deciding on the final disposition of a manuscript, the chief editor may work alone or in consultation with other editors or a board; may or may not feel constrained by referees' judgments; and may or may not work closely with authors of basically sound, but not yet publishable, manuscripts. This will depend to some extent on the age, the prestige, the depth of the manuscript pool, the number of staff, and the sheer size of an annual volume of the particular journal. It will depend even more on the editorial understanding of the journal's purpose—that is, whether it exists to vent all serious work of any potential value or to exclude all but the very best manuscripts, thereby guaranteeing the highest possible quality in what is published and the certain rejection of some worthy but middling work. Rejection rates, which range from less than 10 percent to more than 90 percent, vary in response to all of those factors—but equally influential, and especially interesting, is the impact of a journal's subject matter.

Generally speaking, humanities journals reject the largest proportions of submissions, social sciences the next largest, and hard sciences the smallest. There is, of course, great variation, and a particularly prestigious scientific journal, especially one of fairly broad subject scope, may reject most of the manuscripts it receives. For example, in 1978, The Lancet claimed a rejection rate of 83 percent and Nature a rate of 65 percent, which were remarkably high for scientific journals. However, such rejection rates are standard in even the less distinguished social sciences journals and would be quite low for any humanities journal. In the same year, Economica and Mind were rejecting 90 percent of submissions and Philosophy 92 percent (Gordon, 1978, p. 37).

One obvious reason for these differences is the much larger number and size of scientific journals. But the reason for that is the sciences' different attitude toward research and publication. For anything to remain unpublished if it has the slightest chance of contributing to the advancement of knowledge is anathema—
publication must occur and quickly. This attitude is shared to a considerable extent by social scientists, especially those in the more self-consciously scientific disciplines such as psychology; but it makes much less sense to humanists who tend to be more concerned with arriving at illuminating interpretations than with unearthing facts. John S. Rigden (1986), editor of the *American Journal of Physics*, has pointed out that:

> the humanist brings subjective criteria to the review process. The quality of the writing, the perceived significance of the thesis developed, the inherent interest of the subject, the appropriateness of the context chosen for the subject, and the treatment of the subject-context interaction all influence a recommendation....A review that baldly states, "This looks all right to me," would be unacceptable in the humanities. (p. 491)

Also, of course, timeliness may matter less, and book publishing more, to the humanist. While humanists build on one another's work, the procedure is usually not so linear, the link not so direct, and the related evidence not necessarily so exhaustively assimilated. At the same time, and for the same reasons, absolute factual accuracy and full description of methodology, while expected in the humanities, are far more critical in the sciences, as readers assess the authors' accuracy and often plan costly research projects that build on the authors' work. So, although less is screened in the sciences, the effectiveness of screening may be thought to matter more. Not surprisingly, most critiques of peer review have been written by scientists, with fewer by social scientists, and only an occasional published comment by humanists.

Many of these critics have focused on the question of anonymity in peer review, or as it is typically called, using a peculiarly inapt metaphor—"blindness."

**The "Blindness" Controversy**

As we shall see, peer review is censured for, among other things, its alleged corruption by referees' personal loyalties and biases favoring well-known authors and prestigious institutions. Put another way, it is said to penalize women, minorities, the young, the obscure, and those affiliated with third-string colleges and universities (to say nothing of "independent scholars"). To solve this problem, or simply to forestall any suggestion that it exists, some journals "blind" their referees—that is, conceal from them the identity of manuscripts' authors. When combined with the much more common practice of hiding referees' names from authors, this is called "double-blind" peer review. Both stages of blinding have been questioned.

Referees' names are concealed allegedly to assure that their judgments will not be compromised by reluctance to alienate their
peers or, in the case of younger scholars reviewing the work of those older and better-known, by a natural desire to protect their professional futures. Manfred Kochen (1978) has pointed out that an eminent author who knew his work to have been substantially criticized by his junior might take offense (p. 241). Quite possibly he would dismiss both the criticism and the journal, to no good effect and to everyone's detriment. Indeed, any author able to shift focus from the soundness of the criticism itself to the person behind it may do so.

Jean D. Wilson (1978), of the American Society for Clinical Investigation, has suggested that known referees might be subjected to "face-to-face or telephone encounters with irate authors (p. 1700). See also, Manheim, 1973, pp. 534-35). Michael Gordon (1978, p. 240), and Norval D. Glenn (1976, p. 182) are among those who have warned that, ordered to sign their reviews, some prospective referees would refuse to write any, which could create serious problems for the many journals that find it difficult to attract and retain competent reviewers. Payment for their service ranges from token to none with none prevailing, so the journal peer review system depends upon referees' sense of professional obligation, generosity, and good humor.

However, several commentators insist that, if forced to sign their names, referees would be more thorough and responsible and could be challenged directly by authors, with often fruitful results for both specific manuscripts and general scientific discourse (see, for example, Mirman, 1975, p. 837; Lindley, 1984, p. 59; Raza & Preisler, 1985, pp. 470-71; Nield, 1985, p. 65; Bardach, 1988, pp. 516-17); some have gone so far as to recommend that referees' reports be published alongside the papers in question much more often than is permitted by the occasional symposia seen now (see, for example, Armstrong, 1982, p. 87). Also, authors may be best able to evaluate and profit from criticism if they know the background of the critic (Newman, 1966, p. 980). And if reviews were to bear their signatures, high-status, over-committed referees would presumably be less likely to hand off the chore of writing them to subordinates, unacknowledged (Douglass, 1985, p. 270).

Perhaps the most eloquent opponent of referee anonymity has been scientist-activist Barry Commoner (1978). He sees reviewers' mistakes as equal to authors' errors in their ability to impede scientific progress, and, because they reflect publicly on no one's name, as less likely to be corrected (p. 26). And R. Douglas Wright (1970), an Australian physiologist, demands: "Why should the wish to publish a scientific paper expose one to an assassin more completely protected than members of the...Mafia?" (p. 404). Still, virtually all journals blind authors to referees' identities or at least leave the matter up to the referees if only to avoid offending them and losing their
service. For example, the new *Journal of General Internal Medicine* disapproves of anonymous reviews, yet only encourages—does not require—referees to sign them (Editorial Staff, 1988, pp. 412-14).

Whether authors' names should be concealed from referees is a question less firmly settled in practice, though surveys across a variety of disciplines have found that more journals do not conceal names than do (see, for example, Budd, 1981, pp. 77-81; Miller & Serzan, 1984; pp. 683-84; Weller, 1987, p. 34; Cleary & Alexander, 1988, pp. 1001-02). It may be difficult given authors' tendency to self-cite and drop other identifying clues. And it is often asserted that in very small fields or "cutting-edge" research areas, qualified referees will be able to guess whose work confronts them. However, Moossy and Moossy's (1985) study, which found that referees for the narrowly focused *Journal of Neuropathology and Experimental Neurology* correctly named authors of submitted manuscripts only 34 percent of the time, casts doubt on this "truism" (pp. 225-28). It can also be argued that an author's status and experience are not irrelevant to assessing the authority of his remarks, and it is best that the referee know his name.

The National Enquiry [Committee] into Scholarly Communication (1979), formed with some fanfare over a decade ago, was skeptical of whether benefits accrued from authorial anonymity but concluded that it might be desirable even if it served only to reassure young, or female, or poorly "connected" authors that peer review was fair: "The credibility of the process is of great importance" (p. 48. For an opposing viewpoint, see Evans, 1986, p. 158). This is true, of course, because the development of new knowledge must proceed from what is already known, which is most widely disseminated through journals authenticated by peer review systems. Forward movement of research and analysis requires well-founded faith in these systems: faith that what they include has merit and what they exclude does not.

**The Positive Impact of Peer Review on Intellectual Freedom**

Standing, theoretically, between scholarly editor and author are the author's expert "peers"—though actually few referees are perfect peers, having rather more or less knowledge than the author. Beholden, again in theory, to no one and caring about nothing but the value of the author's work, referees form a defense against carelessness and corruption. Whether or not the editor is obliged to heed their advice, he certainly tends to be influenced by it. Thus referees protect authors from editors—from their whims, biases, and ignorance—and protect readers from both. This is most true, of course, if there is little editorial screening of papers to be refereed (few journals
send out all submissions), if referees are chosen well and objectively, and if they then perform well and objectively. An editor managing submissions alone or with the help of a small staff would soon crash against absolute limits of time and knowledge, with the likely result that, when screening manuscripts, she would seek easy cues such as trendiness of subject, author's status, even her own friendship with the author. D. A. Pyke (1976), a London physician and referee, points out that even within his specialty of diabetes, there are vast areas about which he knows little—which leads him to speculate that the editors of *Diabetes* and *Diabetologia* use referees partly because, despite the narrowness of their journals' scope, they simply do not know enough to evaluate all the manuscripts they receive (p. 117). Geneticist James V. Neel (1988), concerned about the effects of an editor's not knowing what she or invited editorialists do not know, called for the extension of peer review to editorials (p. 981).

Offering, as a group, not only diverse knowledge but wide-ranging backgrounds and alliances, referees are thought to inhibit the development of a "charmed circle" around an editor. Although some repetition of authors published and commonality of style and research approach can enrich a journal, shaping its identity, too small, tight, and dominant a circle of like-thinking scholars may stifle it, thereby closing off fresh ideas and methodologies.

Most obviously, however, referees are employed to prevent readers from being damaged (and editors from being embarrassed) by the dissemination of untruth as fact. The most egregiously harmful results may be seen in professional practice—when, for example, a physician misdiagnoses or mistreats a patient (for examples of harmful medical misinformation, see Knapp, 1988, pp. 371-72; Robin & Burke, 1987, p. 253), or an educator selects the wrong approach to teaching a learning-disabled child. But equally essential is the prevention of futile, costly attempts to replicate faulty research. And, ideally, the peer review process sifts out what would become the trivial, useless, and misleading components of "information overload"—a phenomenon which, in our time of proliferating publication, forms a peculiarly insidious constraint on intellectual freedom. Trying to detect which few items in the onslaught are true and crucial, readers may become captives to an impossible intellectual task, not knowing how to proceed, where to stop, what they are missing, or how or when or whether to act on what they learn. They lack confidence in their ability to access scholarship effectively, and, lacking confidence, cannot assert control, cannot be free.

On the other hand, conscientious peer review may release ideas and information that would otherwise languish in an author's desk, or be published with such severe deficiencies in presentation as to discourage or even mislead most readers. While referees are sometimes
criticized for demanding pointless changes, they are also lauded for helping authors shape poorly written manuscripts with worthy content into readable, persuasive, important journal articles (see, for example, McCartney, 1973, p. 440; Nowell, 1978, p. 844; Bailar & Patterson, 1985, p. 654; Last, 1985, p. 455; Spodick, 1986, p. 3862). Through his survey of 361 statisticians and psychologists, James V. Bradley (1981) discovered much discontent with referees but general approval of their work as advisors on revision. Seventy-two percent of his respondents thought they had improved their writing by following referees' recommendations, while only 5 percent claimed to have had work degraded by referee-induced changes (pp. 32-33).

Peer review, then, is intended to open doors and clear pathways for authors and readers, liberating them from editors' biases and limitations; pre-screening inaccurate assertions, silly interpretations, and data drawn from unsound research; and facilitating the exposure of material that can advance their knowledge. The system ensures that work will be accepted based only on its merit, objectively determined, and guards against decisions influenced by fashion, friendship, and reputation. These, at least, are the justifications for peer review. In practice, quite different things may happen.

NEGATIVE IMPACT OF PEER REVIEW ON INTELLECTUAL FREEDOM

Minor Factors

First, peer review adds days, weeks, or, if poorly managed, even months to the period between manuscript submission and acceptance or rejection. Though publication delays are not the greatest threats to intellectual freedom, they become significant to authors when timeliness of research and the establishment of priority—that is, of "ownership" of a discovery—weigh heavily. The related authorial desire for rapid publication to support bids for promotion, tenure, or grants, would not, in an ideal world where research and writing were pursued for only their inherent satisfactions, affect the journal or, indeed, exist at all. However, the academic world is no closer to ideal than any other, and extreme delays in publication can jettison chances for grant funding for young scholars or even cost them their job; either result will, of course, greatly undermine their freedom to pursue their intellectual interests.

To reader-researchers positioned outside scholarly networks or working in areas tangential to their usual specialties, any delay in publishing the new findings of others may inhibit their progress and place them at a disadvantage relative to colleague-competitors in their fields. Any constraint on access to needed information fetters intellectual freedom. When the constraint is perceived as unnecessary,
it becomes intolerable, and many critics believe that peer review takes far longer than it must. The number and time span of published discussions of this straightforward issue are surprising (see, for example, Newman, 1966, p. 980; Rodman, 1970, pp. 351-57; McCartney & Leavy, 1973, pp. 146, 287-88; Meadows, 1977, pp. 787-93; Azbel, 1978, p. 82; Stieg, 1983, pp. 106-07; Sattelmeyer, 1989, pp. 173-77), as are some time-lag study findings. For example, authors of manuscripts published in *Physical Review* in the early 1980s received acceptances in anywhere from 16 days to 666 days, or nearly two years after submission; mean time lag for all manuscripts accepted by *Physical Review* and *Physical Review Letters* (1979-1980) was 125 days, or more than four months (Dehmer, 1982, p. 96). Brackbill and Korten's (1970) survey of psychologists, taken more than ten years earlier, presented respondents with a list of twenty-two suggestions for revising journal review procedures and asked them to indicate agreement or disagreement with each. Strongest agreement (4.55 on 1-5 scale) was with: "Measures should be taken to insure speedier review of articles" (p. 938). But John Budd (1988) found that seventy-four humanities journals took an average of three months to decide on acceptance or rejection but twelve months from acceptance to publication (p. 183). A second Budd (1988) survey of library and information science journals repeated his earlier finding that producing takes longer than deciding (p. 127). Still, time to acceptance may be especially important to the author, though not to prospective readers. Some journals seek ways to reduce review time—e.g., by computerizing selection of referees and enforcing short deadlines for their reports. However, as long as the number of submitted manuscripts remains at its present almost overwhelming level or (more likely) continues to grow, qualified reviewers will be in short supply and will continue to be so burdened that even the most cooperative may sometimes delay sending reports.

Posing a more serious threat to intellectual freedom when it occurs, but apparently occurring only rarely, is outright referee bias. Though editors may be biased, too, they are carefully screened for their jobs, and they are known and answerable to authors and readers, which tends to make them avoid overt discrimination on bases other than quality. So while peer review may correct for editors' biases as suggested earlier, it is not itself a bias-free process. At journals where editors accept reviewers' recommendations more or less unquestioningly, there may be no corrective for their biases except, of course, other reviewers' opinions.

Often alleged, though rarely if ever proven, is prejudice against women and minorities (few research projects purporting to explore this question have gone beyond simple tabulations of authors' sex. One that did, a study of manuscript reviews for *Rural Sociology,*
"did not indicate that gender was related to editorial decision outcomes...." See, Warner et al., 1985, p. 618). One of the several reasons for "blinding" referees is to avoid this. Partly, perhaps, to disarm criticism of its treatment of women authors, PMLA instituted blind review in 1980 (Showalter, 1984, p. 851). But even this does not, of course, eliminate prejudice against women- or minority-focused research topics, which is sometimes thought to be a greater problem. For example, PMLA has also been assailed for printing few articles about women authors (Gale, 1987, p. 8). To help right both inequities, special efforts have been urged to recruit female and minority referees (APA Committee on Women in Psychology, 1980; Exum, 1983, pp. 127-28).

Also more often claimed than documented is clear-cut political bias—that is, against research or analyses obviously generated by a "left" or "right" viewpoint. Systematic research on this question, as on other questions raised by peer review, is scant. An experiment conducted by Abramowitz, Gomes, and Abramowitz (1975) found that referees, especially liberals, tended to favor work that dovetailed with their political sympathies. However, the evidence was not extremely strong, and the United States of the early 1970s may have been insufficiently typical of other times and places to permit generalization.

Biases favoring the referee's own institution, alma mater, or country have sometimes been hypothesized and would seem to be quite likely outgrowths of natural human weaknesses. But again, direct evidence is lacking, and positive biases stir less ire than negative biases except when competition for page space is very fierce and grossly inferior work is finding print.

Of much more concern are instances of bias describable as idea-based. These arise from intellectual and commercial conflicts of interest that compromise the referee's objectivity and could even extend to "stealing" the author's work: Steven H. Gale (1987) warns, "sometimes experts are the worst people to ask to serve as referees" (p. 12). That is, they almost invariably have strong opinions, sometimes quite emotionally held, about what should be studied and how, which findings and interpretations are plausible, who should be cited, how the writing should be styled, and conversely, what is impermissible. D. A. Pyke (1976) finds it necessary to caution his colleagues "to resist the temptation to advise acceptance of a paper merely because it makes frequent (and favourable) reference to your own work" (p. 1118), though one would expect a sharp-eyed editor not to select a cited authority as a referee. In any case, references to a reviewer's friends, mentors, or co-authors can easily show up and may be seductive to him if the author is approving, or infuriating
if the author is negative. Of course the referee may also be influenced by affection or animosity toward the author himself if he is not “blinded” or is able to guess the author's name.

Physicists Henisch and Roy (1977) point out that in the ever more specialized world of the sciences, few people may be capable of understanding any given piece of work, so: “The chances are overwhelmingly that a submitted paper, handled along traditional lines, would go to a direct professional rival” (p. 105). And a rival can, if he wishes, attempt to torpedo even the strongest argument. In his send-up of peer reviewers' behavior, psychologist Richard Nisbett (1978) writes:

> If the study engages the subjects' interest and deals with matters that are important to them, then assert that the findings were obtained only because of the motivations or defenses that were aroused by the procedures....If the study does not engage the subjects' interest, then so much the better. It may be claimed that the phenomenon under study would not hold up under any but the barren laboratory situations studied....If the instructions to subjects were long and complicated, assert that the subjects probably didn't understand them. The same criticism may be applied if the instructions were brief....any reviewer worth his salt can think up as good a theoretical position as the author's in a few minutes' time. The author may then be criticized for failing to take this position into account.... (pp. 519-20)

In the darkest scenarios painted by peer-review critics, a referee-rival may advocate rejection of sound work either because it disagrees with his preconceptions or to gain an advantage by undercutting the author's career (see, for example, Wilson, 1978, pp. 1699-1700; Wright, 1970, p. 404; Goidon, 1977, pp. 342-43; Oppenheim, 1980, p. 7). In gross ethical violations, the referee may procrastinate or demand trivial revisions in order to delay publication until after the appearance of his own article on the same subject (see, for example, Rodman, 1970, p. 355; Meadows, 1977, p. 791). Or, if involved in related research, he may take unfair advantage of having early access to the author's findings. He may even plagiarize the manuscript, a hilarious happening in Kingsley Amis's college satire, *Lucky Jim*, but deadly serious when it occurs in real life. “Who has not heard of or been the victim of a review in which a view that was...antithetical to the reviewer's preconceived notions or prejudices was suppressed merely by giving the article a bad review?” Harry C. Nottebart, Jr. (1982), a physician, once demanded. “Who does not know of situations in which a reviewer has used data from someone else's work that was being reviewed” (p. 480)? Finally, and perhaps most egregiously, a referee may turn the still-confidential information contained in a manuscript to immediate financial advantage, an increasing possibility in areas like biotechnology where many researchers have undisclosed links with business (see Maddox, 1984, p. 497; Vevaina, 1987, p. 958).
All of these potential problems, however, from sex and race discrimination to conflict of interest, are reassuringly straightforward and probably very infrequent compared to the more subtle and pervasive controls imposed by peer review. Before we consider those, however, a simpler issue presents itself: are valid, reliable judgments even possible under peer review?

**Major Factors**

In any gathering of scholars where the conversation turns to peer review, stories will be told of incompetent reviews—careless, cursory, uninformed, uninformative—used to assault manuscripts representing months or even years of painstaking work (for examples of such stories, see Lindley, 1984, pp. 56-58; Commoner, 1978, p. 25; Wright, 1970, pp. 403-04; Glenn, 1976, pp. 179-80; Raza & Preisler, 1985, p. 470; Sommers, 1983, p. 92; Engler et al., 1987. For a pointed, amusing satire of referees’ tactics, see, Remus, 1980). Astute reviews, which are rarely discussed, are probably more numerous. But the fact remains that, even allowing for authorial egotism, some very poor judgments seem to be made by referees and presumably cause at least occasional rejections. According to Robin and Burke (1987): “For some [medical] journals, even one unfavorable review may diminish the priority and result in disapproval” (p. 254). In James V. Bradley’s (1981) aforementioned survey of statisticians and psychologists, 74 percent asserted that for the most recent of their articles published in a refereed journal after compulsory revision, at least some factual errors were made by the referees; 42 percent of all respondents had found errors in important facts while another 32 percent found only trivial errors. In addition, 67 percent claimed that at least some of the referees appeared not “to be at least as competent and sophisticated” in the article’s subject area as they, the authors, were; 40 percent said that at least some referees did not seem “to have read the article carefully” (p. 32).

Editors try to avoid such problems by inviting high-status scholars to serve as referees in the belief that they know the most and will render the best reports. Thomas P. Stossel (1985), editor of the *Journal of Clinical Investigation*, carried out a fascinating experiment that stands this assumption on its head. Stossel counted review requests, analyzed the professional status of those to whom requests were sent, and evaluated the quality of completed reviews. He found that the highest-status scientists were the most likely to refuse his requests and, when they did comply, were the most likely to provide low-quality reviews. Conversely, the lowest-status scientists were likeliest to grant his requests and to write high-quality reviews.
His conclusions raise some alarm as high-status people routinely referee and their reports are routinely accepted and not scrutinized and assessed.

Striking signs that something is amiss with peer review are the low levels of agreement among referees and, after publication, between referees and readers. That is, just about as often as not, those refereeing the same paper make different recommendations, and there seems to be no correlation between strength of referee endorsement and numbers of times cited after publication. In their survey of 138 refereed education journals, Smith and Gough (1984) asked editors to estimate "the percentage of manuscripts that provoked significant disagreement among the referees" and received answers ranging from zero to 90 percent with seventeen declaring that their referees rarely agreed (pp. 638-39).

Nor are the social sciences unique. For instance, biological scientist-referees are said to show a rate of agreement about the same as would be reached through chance alone (Wilson, 1978, p. 1698). William C. Roberts (1987), editor of the American Journal of Cardiology, who uses two referees per submission, observes that it is unusual for both to write "definitely accept" or "definitely reject," and occasionally, perplexingly, one will write the first recommendation and one the second (p. 922). In the physical sciences, referees are much more consistent, presumably because their methodologies and rules of evidence are more firmly fixed, their research material by its nature more stable, and true objectivity is more easily attainable. However, even physicists and mathematicians tell of startling conflicts among reviewers (e.g., Lindley, 1984, pp. 57-58; Wallace, 1983, pp. 11, 13; Thompson, 1983).

In what is probably the most famous and controversial study ever conducted of peer review reliability, Douglas P. Peters and Stephen J. Ceci (1982) randomly selected one prestigiously-authored article published recently in each of twelve major psychology journals. They then changed the title and author's name (but not gender) of each article, invented a low-prestige institutional affiliation, altered the abstract slightly, retyped the article, and submitted it to the journal where it had appeared. All were journals that made it a practice to reveal authors' names to reviewers. In only three cases did any editor or referee recognize the article as previously published, and in eight of the remaining nine cases, the article was rejected. This suggests, of course, not only stunning unreliability in editorial decision-making, but the failure of referees to keep up with the literature, the failure of editors to read even their own journals, and probably institution-based discrimination—though it is unclear whether authors from low-prestige institutions were being wrongly penalized or authors from high-prestige institutions wrongly favored.
Controversy swirled around the method, ethics, and significance of the Peters-Ceci study, but at the very least it raised questions about peer review that were difficult to dismiss.

Two reasons are most often hypothesized to explain the system’s weaknesses—unmanageably heavy demand for referees and lack of concrete guidelines. Stossel’s (1985) study mentioned earlier, correlating reviewer status with review quality, suggests the first of these. The numbers of journals, journal pages, monographs, and grant proposals seem to have been growing faster than the number of experts with ample time to referee. A few of these numbers suffice to sketch the problem: Fifteen years ago, *American Sociological Review* was receiving about 800 manuscripts each year, *American Journal of Sociology* about 700, and together they needed 3,000 referees’ reports in order to dispose of the load (Glenn, 1976, p. 180). Eleven years ago, *Physical Review Letters* was soliciting approximately 8,000 referees’ reports each year (Adair, 1979, p. 101). In 1985, the *Journal of the American Medical Association* received 3,446 manuscripts, all of which were screened by the editors, 1,413 of which were refereed (Lundberg & Carney, 1986, p. 3286). Five years ago, Jay H. Lehr of *Ground Water* assured his readers that the journal’s referees were not overburdened: “No referee receives more than two papers a month [or] more than eighteen papers a year” (p. 148). One assumes that the readers were not reassured. It seems to take close to a full working day for thorough review of an exacting paper (see, for example, Carney & Lundberg, 1987, p. 87; Curtis, 1982, p. 9), and capable reviewers may have their opinion solicited regularly by more than one journal.

On first consideration, the lack of written standards to serve as review guidelines seems much easier to rectify, and indeed many journals do supply them. However, they are hard to win agreement upon; hard to word specifically yet flexibly enough to cover all submissions; hard to formulate so as to yield valid judgments; hard to enforce; and almost certain to be variously interpreted (for examples of guidelines, see, Forscher, 1980, pp. 166-67; Bishop, 1984, pp. 59-67).

If, as seems true, substantial numbers of peer reviews are compromised by prejudice, ignorance, carelessness, hurry, or uncertainty or misapprehension about the journal’s values, many authors and many more readers are being arbitrarily denied opportunities to be heard and to learn. This, however, is not the greatest threat of peer review to intellectual freedom. Even when the system seems to work smoothly—perhaps especially when it does—it may subtly and harmfully control not only what is published and read but what phenomena are investigated and what ideas pondered.
RESISTANCE TO THE NEW AND UNCONVENTIONAL

The peer review process is inherently conservative because of the way reviewers are selected. Chosen by, and answerable to, an editor or editorial board (sometimes after consultation with colleagues or other reviewers), they are most often scholars with research degrees, affiliations with well-known academic or research organizations, and publishing histories of their own. The more famous and established they are, the more likely they are to be asked for reviews. Zuckerman and Merton (1971) discovered, for example, that: “Relative to their numbers, lower-ranking physicists do little refereeing altogether and also referee far fewer papers by the intermediate and highest ranking physicists than would be the case under a populistic allocation” (p. 89). This is understandable, of course, but it guarantees that the vast majority of reviewers will have many years’ experience as students and employees of mainstream learned institutions, and probably as writers for refereed journals and as successful grant applicants. They will have absorbed the associated values and norms, acquired distinct theoretical and political perspectives, and formed opinions about the appropriate credentials for researchers and directions for research—they will have become, in other words, what Peter Gibson (1987) has called “the Enid Blytons of the scientific fraternity” (p. 63). However open-minded they strive to be, their judgments are bound to be shaped by powerful past influences and present expectations. That is, they not only have certain backgrounds but share them with those who request and receive the reviews, because scholarly editors are drawn from the same general population (though perhaps slightly more elevated ranks of it) as reviewers are (on referee selection, see, Stieg, 1983, pp. 102-05; Gordon, 1978). An interesting assertion was made by the editor of Journalism Educator when he initiated a refereeing system: “Manuscript reviewers will be chosen from the ranks of established scholars and professors who are likely to offer innovative ideas about the field” (Crook, 1988, p. 56).

When consensus among reviewers, or even a majority “vote,” is required for acceptance of a manuscript, the tendency toward safe, unexceptionable decisions and avoidance of intellectual risk-taking is likely to be especially marked. And in high-rejection journals where, to use Zuckerman and Merton’s (1971) terms, the “decision-rule” is “when in doubt, reject” (p. 78), adventuresome manuscripts must be anathema. As a result, unconventional subjects and ideas, novel research designs, findings that challenge long-standing beliefs, and anything controversial—even when persuasively presented—may find the road to print rough and darkened by shadowy obstacles. “Novelty,” states Dennis V. Lindley (1984), “is always on dangerous ground with referees” (p. 57). Physicians Robin and Burke (1987) observe: “Almost
every expert reviewer has a conflict of interest; he or she represents their discipline as it now exists and unconsciously tends to defend it” (p. 254). And Barry Commoner (1978) warns:

The peer review system threatens not so much the bulk of more routine research, but precisely those sweeping, novel advances that are the growing points of science. The real danger in the system is that it threatens to blunt the cutting edge of scientific progress. (p. 29)

Also, of course, this blunts progress and path-breaking exploration in the social sciences and humanities. Convincing research on the subject is difficult to plan and carry out since so many variables may affect what is finally a qualitative decision, but some investigators have made the attempt with provocative results. For example, Zuckerman and Merton (1971) found that high-status physicists were likelier than their intermediate- and low-status colleagues to have manuscripts submitted to Physical Review accepted at all, to have them accepted immediately, and less likely to have them rejected immediately (p. 91. See also, Crane, 1967). Of course, high-status authors may simply do better work than others, hence their high status.

Steven H. Gale (1987), a literary critic with an axe to grind, subject analyzed twenty-two consecutive issues (1978-83) of the Publications of the Modern Language Association (PMLA), containing 146 scholarly articles, and concluded that, “the referees tend to look for the same kind of material that has already appeared in the journal...” (pp. 4-9). (It is not irrelevant that Gale’s specialty is the contemporary English dramatist Harold Pinter and that his Pinter articles were rejected by PMLA). There were, for example, disproportionate numbers of articles on pre-twentieth century English literature, on French literature, and specifically on Shakespeare and Chaucer. Largely neglected were drama and American literature—and only eight articles dealt with women writers of any country or period (Gale, 1987). The issue is not clear cut, of course: in an earlier PMLA editorial, English Showalter (1984) had emphasized that what he published simply reflected what was submitted (for example, 46 percent of manuscripts received in 1973-83 dealt with English literature) (pp. 851-53)—though it does seem odd that, in the period Gale studied, so few manuscripts about women’s writing would have been submitted to one of the field’s most prestigious and broadly subject-defined journals.

Michael Mahoney attracted considerable attention with his studies of social science journals which showed that experimental data and conclusions that supported conventional theory were much likelier to win referees’ approval than those that conflicted with it—and that articles citing work “in press” from the author were more often accepted than those that did not. “Even in science,” Mahoney
(Quoted in Joyce & Pearce, 1986, p. 23) lamented, "nothing succeeds like success. Publication begets publication, recognition begets further recognition, and the rich get richer" (p. 23). It is the phenomenon that Robert K. Merton (1973) famously labeled "the Matthew effect."

In the American Council of Learned Societies' 1985 survey of 3,835 humanists and social scientists, 77 percent of the respondents asserted that the "peer review refereeing" systems in their fields were at least sometimes biased in favor of "scholars who use currently fashionable approaches" and "established researchers in a scholarly specialty." Seventy-three percent believed that "scholars from prestigious institutions" were often or sometimes favored, and 37 percent felt that "males" were favored. Women were likelier than men to detect "frequent" bias of any type (Morton & Price, 1989, pp. 69-71). (What was not asked was whether respondents saw the biases as justified. In a follow-up study of educators in ALA-accredited programs, 67 percent, 71 percent, and 67 percent of respondents perceived at least some bias favoring authors in the first three categories respectively, but several added marginal notes to the effect that discrimination in favor of the established and prestigiously employed makes sense [Biggs & Biggs, 1990].)

Suffering most in such circumstances will be younger scholars without influential mentors, substantial bibliographies, or impressive institutional ties; writers in developing fields and new interdisciplinary specialties; intellectual innovators and rebels of all stripes; and, some suggest, authors reporting applied or apparently "simple" research. (See, for example, Grassman's [1986] amusing critique of mathematical journals. Based on rejections of his articles on queueing theory, Grassman formulated "Joe's theorem": "Nothing is published in the area of queueing theory unless it is mathematically interesting. Nothing is applied in industry unless it is mathematically trivial. Since trivial results are not interesting, and since results that cannot be applied are not useful, nothing useful will ever be published in queueing theory" [p. 44].)

Some may respond that there is no problem, that most papers rejected by prestigious journals will still find print, if their authors are persistent, through less-known, perhaps unrefereed, outlets. This is true (see, for example, Stieg, 1983, p. 115; Yankauer, 1982, p. 239 [footnote]; Slater, 1984, p. 455; Rennie, 1986, pp. 2391-92). But it may be that in this age of bewilderingly prolific publication, to appear in an obscure journal, especially one that is not refereed, is to remain invisible and not really to "appear" at all. Unable to read every journal in their fields, scholars seek external clues to aid selection; among the most prominent are the related factors of a journal's reputation and whether it is peer reviewed. Peer review is assumed to ensure
adherence to scholarly standards. When it doesn't work well, readers are misled, as are academic search committees and tenure and promotion review groups. And authors who publish in unrefereed journals may suffer wrongly by comparison with their counterparts who publish in refereed, but not better, journals.

**Impact of Peer Review on Intellectual Freedom:**

**Summary**

Through referee procrastination, carelessness, ignorance, or cupidity, peer review may deny authors the chance to publish in the most appropriate and widely read journals. And when authors' names are revealed to referees or when referees can guess them, such essentially extraneous factors as authorial fame (or lack thereof), institutional affiliation, and sex may affect publication decisions.

Even the most vocal defenders of peer review concede that it penalizes innovation and nonconformity. Of course, this harms individual scholars, authors, and readers who are interested in the new. More broadly, it retards the advancement of knowledge, not only impeding progress but undermining hope of progress. For thoughts that cannot be voiced will less often be thought; subjects that cannot be published will virtually cease to be explored; and research approaches scorned will be abandoned. Self-censorship is necessary for the scholar wishing to succeed in academe. That this is so can largely be laid to the account of the peer review system.

**Conclusions and Recommendations**

But what would we do without peer review? Unrestrained publication of everything, which is quite possible in the computer age and is advocated by some, would probably result in true publication of little. That is, as the flood of supposedly available information increased, ever less would actually be accessed and read. And a prestige ranking would surely emerge—probably based on criteria even less valid than presence of peer review—because scholars would have to find some way to differentiate among publications in order to choose which to read and how to assess colleagues' credentials.

In most unrefereed periodicals, manuscripts are screened but by an editor rather than referees. To substitute unaided editorial judgments for referee-assisted judgments in scholarly journals seems not an improvement and essentially absurd in an age of extreme specialization. Another possibility is to limit reviewing responsibilities to an editorial board of scholars in necessary specialties; they would be compensated, their names known, and their expertise and conscientiousness proven. At many journals, however, they would soon be overwhelmed by the volume of submissions, and the inherent
conservatism of peer review would only be enhanced if a small circle of people took charge of all decisions. Alternatively, such a board might divide responsibility, according to specialty, for screening manuscripts and selecting reviewers. Each specialist would then carefully read and evaluate the reviewers' reports and make the decision to accept, reject, or request revisions. Constructive suggestions drawn from the reports could be conveyed to authors, with editors committing themselves to helping authors improve their manuscripts.

Reviewers' competence and care are crucially important and often criticized. Computerized referee files, which now are easy to set up, allow editors to store and retrieve large amounts of information on thousands of potential referees, thus drawing on the opinions of a far wider range of people than personal acquaintances and "invisible college" referrals can provide. Also, their performance can be monitored closely, with careless, uninformed, and dilatory referees identified at once and removed from file. Some editors have created huge databases of prospective referees through questionnaires and literature searches. Editorial involvement is the key: to selecting good referees, using their reports well, and guiding and controlling the entire process.

Guidelines for peer review are probably necessary, though they are tricky to write and may shape and constrict referees' thinking, thereby closing off spontaneous reactions and novel ideas. Certainly deadlines should be imposed on reviewers and self-imposed on editors. Though judging the authority of an anonymous manuscript can be difficult, the arguments for "blinding" referees are persuasive. And while people may indeed devote more care to reports that bear their names, revealing referees' identities seems likelier to cause problems than to solve them. Younger and more vulnerable scholars, in particular, would either decline invitations to review or avoid harsh judgments of their seniors. This would only strengthen the grip of established scholars and ideas. Intelligent, sensitive editorial use of referees' reports would circumvent most of the problems associated with reviewer anonymity.

Among interesting possibilities are to encourage voluntary signing of reviews; to ask that authors suggest appropriate referees for the articles they submit as well as referees to avoid; and to publish symposium-style, immediately following an article, any particularly insightful referees' reports (with their writers' permission).

Finally, though, editors and readers should realize that peer review is fundamentally hostile to intellectual invention and rebellion. It is the price we pay for reliance on established expertise—a necessary price, but a high one.
REFERENCES


Gale, S. H. (1987). Blackwoods would; PMLA won't; or how to write a PMLA article. (ERIC Document Reproduction Service No. ED 285 148)


Grassman, W. K. (1986). Is the fact that the emperor wears no clothes a subject worthy of publication? Interfaces, 6(March-April), 43-51.


Lehr, J. H. Conference proceedings vs. the refereed journal. Ground Water, 24(2), 148.


