Chapter 10
Peer Review: From "Sacred Ideals" to "Profane Realities"

David R. Johnson and Joseph C. Hermanowicz

Introduction

All modern social systems include people who are formally designated with the authority of evaluating the performance of others in a given system (Zuckerman & Merton, 1971). We thus think of the relationship between managers and employees, teachers and students, coaches and athletes, parents and children, CEOs and governing boards, politicians and an electorate. The same is true in the system of modern scholarship and science. Academic peer review consists of socially structured processes for evaluating scholarly and scientific performance. The faculty members of universities, colleges, centers, and institutes are asked to serve in a role of judge over the quality, quantity, creativity, and originality of a fellow member's performance, constituting a "review by peers."

While "peers" are members of academic communities, and while they are normally members of the same field or work area as the individual whose performance is under review, they are otherwise loosely defined. In relation to the person whose performance is assessed, peer reviewers may be of a lower, higher, or comparable status, employed at the same or different institution, methodologically and/or theoretically similar or dissimilar, older or younger, more or less experienced, "blind" or "known." These permutations arise as a function of the specific type of performance under evaluation and as a result of deliberate and undeliberate decisions in the selection of reviewers by those in charge of overseeing the review.

D.R. Johnson (✉)
College of Education, The University of Nevada – Reno, 89557 Reno, NV, USA
e-mail: drj@unr.edu

J.C. Hermanowicz
Department of Sociology, The University of Georgia, 30602-1611 Athens, GA, USA
e-mail: jch1@uga.edu

M.B. Paulsen (ed.), *Higher Education: Handbook of Theory and Research*,
Higher Education: Handbook of Theory and Research 32,
DOI 10.1007/978-3-319-48983-4_10

Peer review in tenure and promotion assessments, for example, includes a vote from all other tenured faculty members in the unit in which the candidate holds an appointment. All of these members are "known" to the candidate, but they may vary in all of the other stated ways. Assessments of tenure and promotion now often also include "external" reviews from "peers" outside the institution in which the candidate holds an appointment. Those who oversee the review—department chairs or heads—commonly possess considerable latitude to whom dossiers are sent; these external reviewers may vary in relation to the person under review in any number of the above stated ways. By contrast, peer review of articles more often than not transpires as a "blind" process (the identity of the reviewer is not known to the person whose work is reviewed) or as a "double-blind" process (neither the identity of the reviewer nor that of the person whose work is reviewed are known to one another.) Yet even these procedures vary among journals and across disciplines. Single-blind review is prevalent in the life sciences, physical sciences, and engineering, for example, while double-blind review is more commonly found in the humanities and social sciences (Ware & Monkman, 2008). Moreover, the editors overseeing the review process at journals exercise very great latitude in deciding whom to send manuscripts.

Aside from their correspondence to different types of performance, these permutations are also a source of bias in peer review (Lee, Sugimoto, Zhang, & Cronin, 2013; Shatz, 2004), a subject that will assume considerable prominence in the present chapter. Bias arises both as a function of the decisions made by those in charge of the review about which peer reviewers to select, and as a function of peer reviewers themselves who apply "particularistic," functionally irrelevant, as well as aesthetic, criteria to formulating their judgments, rather than universalistic criteria.

Nevertheless, "peers," as opposed to clearly specified super- or sub-ordinates as found in other social systems, are used in the review of scholarship and science because of a belief that professions, of which scholarship and science is the prototype (Gustin 1973), require an expertise from fellow members to form prudent judgments about specialized work (Bess, 1988; Waters, 1989). Peer review thus operates as a key mechanism by which professions are, in principle, self-regulating.

Thus conceived, peer review has been called "the linchpin of science" (Ziman, 1968 p. 148, quoted in Fox, 1994, p. 299): modern scholarship does not operate without it. Given the centrality of peer review to academic work, it is important to take stock of what we know about its aims and ideals, its social organization, and the problems in its actual operation. This is the purpose of the present chapter. Higher education researchers have paid considerable attention to other contexts of peer review—tenure and promotion, post-tenure review, grants, fellowships, and salaries—inadvertently eclipsing the centrality and significance of peer review of publication in higher education.[1] For this reason, this chapter focuses on peer review in the publication process.

[1] For other, secondary arenas where peer review operates, readers can consult illustrative treatments: for example, tenure and promotion (Fairweather, 2002; Hearn & Anderson, 2002; Lawrence, Celis, & Ott, 2014; Long, Allison, & McGinnis, 1993; Perna, 2001, 2005; Youn & Price, 2009);

The chapter has three parts. First, we consider the functional aims and ideals of peer review, its historical origins, and its conceptualization. Special attention is paid to the relationship between peer review, the communication of science, and the reward system in science and scholarship. Second, we examine the social organizational basis of peer review. Specifically, we focus on the expansion and differentiation of peer review, as well as on how peer review can be understood both as a labor process and as a process of "invisible interaction." Finally, in a section on dysfunctions, we examine a substantial segment of work on reliability and bias in peer review processes. The discussion underscores structural and cultural characteristics of peer review that seemingly shatter its idealized image. We conclude the chapter by discussing gaps and limitations in extant research to thereby identify promising paths for future empirical studies.

The Social Aims and Ideals of Peer Review

The specific performance of the scholar or scientist to be judged takes multiple forms. The objects of judgment include research and grant proposals, dossiers for tenure and promotion, credentials for select administrative appointments, teaching, salary, "job talks" and portfolios of faculty applicants, records for the conferral of special awards and honors, as well as, occasionally, records of behavior for the meting out of discipline.

While peer review has several referents, the staple of its usage is found in its operation in the process of publication. The publication process is typically viewed as undergirded by a norm of "organized skepticism," or the social arrangements established to ensure that all scientific and scholarly contributions undergo a fair and proper process of peer-based evaluation prior to becoming a part of certified knowledge (Merton, 1973a, 277–278). This is so because science and scholarship cannot advance determinedly, and thus fulfill its institutional goal of extending knowledge (Merton, 1973a, 1973b), in the absence of work that has been certified as a contribution and placed in the public domain (Ziman, 1968). In short, publication enables *communication*. The communal character of peer review transforms a manuscript, long or short, into consensual "knowledge" (Chubin and Hackett, p. 84). Publication thus becomes the central social process of "constituting" science (Fox, 1994, p. 299).

What is more, all other roles performed by the modern professional scholar and scientist (as opposed to the pre-modern amateur) are dependent on publication (Shils, 1997). Across the gamut of what can be taught in a classroom or virtual medium, to what can be shared through varieties of service, academic roles are

post-tenure review (Aper & Fry, 2003; O'Meara, 2004; Patriquin et al., 2003; Wood & Johnsrud, 2005); grants (General Accounting Office, 1994; Gillespie, Chubin, & Kurzon 1985; Langfeldt, 2001; Liebert, 1976; Roy, 1985); research fellowships (Bornmann & Daniel, 2005; Lamont, 2009); and salaries (Perna, 2003).

dependent upon knowledge that is sanctioned thusly and communicated in a format for others to access (Merton & Zuckerman, 1973). The standardization of publication formats, in both the article and book genres, was not conceived at once, but rather an outcome of evolved practice. This standardization is thus itself social testimony to a functional imperative of sharing science and scholarship (Zuckerman & Merton, 1971). The proliferation of journals (Hermanowicz, 2016a, 2016b; Jacobs, 2013) and the institutional intensification of pressure on faculty to publish (Hermanowicz, 2009) are related, developing phenomena that point not only to the centrality, but also to a contemporary spread and ubiquity of peer review.

As a regulative mechanism in the process of publication, peer review is used by members of academic communities to assess the acceptability of manuscripts and, when judged acceptable, to certify that the work constitutes a contribution to the stock of knowledge. In the case of scholarly and scientific outlets of publication, peer review for a given submission is conducted—in present times—by fellow members of a field, usually numbering between one and four depending on the field and on the specific journal or book press, in conjunction with an editor of the particular outlet. Together these actors operate as the judges who manage standards of publication in the academic system. While representative of the peer review process in operation today, and thus highly recognizable to all academics and indeed a focal concern of many of their livelihoods, these conventions took considerable time to develop. They were, for significant spans of time, not customary to, or indicative of, academic work.

The practice of peer review in the academic publication process is, in the context of organized learning, relatively new. As Zuckerman and Merton (1971) have elaborated, it dates to the seventeenth century—an historical period in which the scientific role was legitimized (Ben-David, 1965)—with the advent of scientific societies and academies.

> The new scientific societies and academies of the seventeenth century were crucial for the social invention of the scientific journal which began to take an enlarged place in the system of written scientific interchange which had hitherto been limited to letters, tracts, and books. These organizations provided the structure of authority which transformed the mere *printing* of scientific work into its *publication* (Zuckerman & Merton, 1971, p. 68, original emphasis).

Two journals emerged close in time to each other in 1665: the *Journal des Scavans* and the *Philosophical Transactions* of the Royal Society. The use of reviewers in these and subsequent journals, however, emerged only slowly and unevenly. Manuscripts to fill journal issues often had to be solicited by the editor, and were often "reviewed" or examined solely by the editor, later perhaps by a member of what came to constitute a small in-house staff, and latter still by a member of a journal's editorial board, though even this specific practice was not always observed. Burnham (1990) has explained how the practice of editorial peer review did not become general until after World War II. Consistent with this historical conclusion, Fox (1994) explained that James McKeen Cattell edited the renowned journal *Science* for the 50 years spanning 1894–1945, used his son to review papers, and only after his own death did the American Association for the Advancement of

Science assume control of the journal and adopt external peer review as a standard procedure.

The explanation of this history is found partly in the twin forces of specialization and quality (Burnham, 1990; Weller, 2001). As science developed, scientists specialized in their research and publication. What is more, the number of scientists grew. The number of scientists grew at dramatic levels in conjunction with the rise and expansion of universities. This pattern includes the emergence of the research university in the United States in the late 1800s, which exerted both national and global effects on an intensification of research (Geiger, 1986). Moreover, the elaboration of the research university facilitated specialization through the creation of academic departments (Shils, 1997). Thus, the net research and publication activity of scientists ascended and began to climb markedly in the early twentieth century.

Consequently, the editorial role of a journal reached a point where editors could not alone assess the merits of specialized submissions and could not alone sort stronger from weaker articles in the growing batch of submissions. "Casual referring out of articles on an individual basis may have occurred at any time beginning in the early to mid-nineteenth century, but *institutionalization* of the process in various editorial settings took place mostly in the twentieth century—either to handle new problems in the numbers of articles submitted or to meet the demands of an increasingly specialized world" (Burnham, 1990, p. 1327, original emphasis).

But while specialization and quality account partly for the institutionalization of peer review in academic publication, they in turn are premised on a condition that, first, precedes their occurrence *historically* and, second, more fully informs the motivations and functional operation of peer review *contemporaneously*. This concerns the *incentive* for scientists and scholars to disclose their newly-found knowledge. The advent of printing provided a technological means to communicate knowledge (Bazerman, 1988, especially pp. 128–150). But, as Zuckerman and Merton (1971) again explain, scientists still placed a premium on secrecy so that others could not steal and appropriate what they had discovered.

Other institutional practices had to arrive to encourage a shift from "motivated secrecy" to "motivated public disclosure" (Zuckerman & Merton, 1971, pp. 69–70). A key such practice was the seemingly simple but dramatically consequential act of the Royal Society recording the *date* on which scientific communications were first received. Henry Oldenburg, one of two secretaries of the Royal Society, oversaw the *Philosophical Transactions* beginning in 1665 and, in handling much of its early correspondence with scientists, acted as an "editor," though no such designation had been made nor had any outline of an editorial role been construed (Zuckerman & Merton, 1971, pp. 68–69).

What was the significance of assigning a date? Writing to Robert Boyle, Oldenburg concurred:

> The society alwayes intended, and, I think, hath practised hitherto, what you recommend concerning ye registering of ye time, when any Observation or Expt is first mentioned… [the Royal Society] have declared it again, yt is should be punctually observed: in regard of wch…hath been written to, to communicate freely to ye Society, what new discoveries he maketh, or wt new Expts he tryeth, the Society being very carefull of registering as well the

> person and time of any new matter, imparted to ym, as the matter itselfe; whereby the honor of ye invention will be inviolably preserved to all posterity (Hall & Hall, 1966, p. 319; quoted in Zuckerman & Merton, 1971, p. 70).

The practice officially established *priority of discovery*, the social function of recording date of publication (Merton, 1973b; Zuckerman & Merton, 1971). The formal authority of an esteemed organization sanctioned when and to whom credit was bestowed for contributions to knowledge. In addition, by recording contributions by date alongside people's names, the practice forwarded the idea of permanence of record, a furtherance of honor of its own accord. Permanence enables a more formal basis of scientific and scholarly archives, which constitute a genealogy of contributions and testifies to scientists' achievements in historical time (Chubin & Hackett, 1990; Zuckerman & Merton, 1971).

The practice of registering the date of contributions received, repeated over time, resulted in a convention in and among the people of science. The convention in turn facilitated the emergence of a norm said to characterize scientists' research behavior: the norm of "communalism," which prescribes the sharing and proscribes the secrecy and withholding of scientific findings (Merton, 1973a). That is, in exchange for sharing their findings and placing them in a public domain, scientists were to be given credit, in the form of recognition, through the systematic process of registering who made a contribution and at what time. Recognition was bestowed at the time by the institutional authority of the Royal Society and, as these processes matured, informally and formally by fellow scientists and the organizations of science and scholarship to which they belonged or with which they were otherwise associated (e.g., departments, colleges, and universities; professional associations, societies, and so on).

The quest for recognition thus itself becomes institutionalized (Merton, 1973b; also Hermanowicz, 1998). The desire for recognition, socially manufactured and maintained, is hitherto an expected behavioral pattern in the life of (effectively socialized) modern scientists and scholars. So conceived, recognition takes its place in the functional process of science. Recognition serves as social testimony from one's peers—those capable of judging scientific and scholarly contributions—that one has indeed satisfied the institutional goal of science, to extend knowledge (Merton, 1973b). "Writing for publication" is a behavior that not only emerges and becomes standardized (Zuckerman & Merton, 1971), but a behavior that comes to be idealized among scientists, for publication serves as the basis of honoring one's own and others' work and, correspondingly, the "ways" of doing science. Great recognition bestowed on scientists implies, by this perspective, a high degree of originality in their work, and thus a proportional furtherance of socially certified knowledge. Through these ways, the force of specialization and the desire for quality, discussed above, do not alone account for the development of peer review. Antecedents are found in the emergent *system* of science itself which, as it modernizes, creates needs and uses of peers to form its very "operational machinery."

For a long time unknown to science, peer review is now so pervasive it signs a signature of legitimacy to the work of scholars and scientists. Editors—and all others who are called upon to review the various performances of academics—can

turn to the advice of peer reviewers to justify their decisions, and to the fact that given work was peer reviewed, to underscore the legitimizing authority that peer review possesses. Tenure and promotion committees, at least in many universities, discount work of candidates that "was not peer reviewed" as well as articles and books that were not subject to "rigorous" (meaning fully genuine) peer review.

Put in starker terms, contemporary academic work that is not subjected to peer review is not even considered scientific; it is "not good." The process of peer review is now wedded to the system of science and scholarship. As Ripp has stated:

> [Peer review] has become part of the reward system of science, and it is legitimated by the ideology that peer review is the best way to conduct evaluation processes. Indeed, questioning the functioning of the peer review system is questioning the ideology, and thus inviting hostile reactions (Ripp, 1985, p. 83).

And, as Campanario has added, illustrating the historical distance that the practices of science have traveled:

> Not surprisingly, publication in academic journals is judged as an indicator of performance of an individual effort, and on many occasions it fosters an author's career advancement… This prestige—and the reason we trust journal quality—rests on the process by which manuscripts are evaluated before publication; that is, the peer review system (Campanario, 1998, pp. 181–182).

It might seem logical, then, that few would dare to question, or indeed criticize, the practice of peer review since, as the modern-day "linchpin of science," it is its guarantor of authenticity. That is, however, precisely what people have done, and their own work forms a body of literature that has developed and grown over the past three decades. Such scholars assert that peer review "seldom operates according to strictly meritocratic criteria, that it is frequently unreliable (in that reviewers, quite predictably, disagree about the merits of the work under review), and that it may be sensitive to influences other than 'true' scientific quality (such as the prestige or past performance of the scientist whose work is under review)" (Chubin & Hackett, 1990, p. 125). By this light, peer review may be viewed not only as a "linchpin of science" but also a "lightning rod" for controversy among scientists.

While central to science, peer review is anything but "pure." The truism is that it involves other human beings, as opposed to insentient mechanical robots. The "performance" of peer *reviewers* themselves vary, not merely by the types of performances that they are called upon to judge, but by qualitative and quantitative terms of their participation in review processes. The issue prompts key concerns: the raw time faculty members spend in service of peer review, the diligence with which they assume such roles, the sources of variation in the quantity and quality of participation in review processes, and the consequences for the institution of science and scholarship of such variation.

On the one hand, peer review is enshrouded in myth and ceremony: it projects what we shall call "sacred ideals." We may say that the ideals of peer review are sacred because they encompass the norms of disinterestedness and universalism, are thus "objective," and consequently assure a positivistic conduct of science (Merton, 1973a). The norm of disinterestedness holds that scientists' performance of scien-

tific work, including their conduct in evaluation, should be free from biases (Merton, 1973a), while the norm of universalism stipulates that the allocation of rewards for achievement should be based exclusively on objective standards (Merton, 1973a).

By contrast, a growing literature demonstrates the existence of what we call peer review's "profane reality." We may say that the reality of peer review is profane because of limitations and failings of peers as human actors, for whom an objectivity is unachievable, which results in a "social construction of science." By this alternative view, science cannot be "disinterested" or "universal"; these so-called "norms" are in fact not indicative of most actual behavior, but rather they are of a variety of ideals associated with a rhetoric and ideology of science, the manifestations of which are highly contingent on context (Mulkay, 1980).

Consider the following statement:

> The work of a scientist should be viewed by disinterested, objective standards. The most appropriate person to judge those standards is an appropriate peer—another scientist. Extraneous and subjective issues such as rank, jealously, affiliation and rivalry should play no place in this process, neither should lack of expertise. This is the theory behind the practice of peer review, the process by which papers submitted to professional journals or grant applications made to funding bodies are usually judged (Wessely, 1996).

And compare with:

> … a high volume of submissions, turnover of editors, and dependence upon ad hoc reviewers and often a part-time (non-professional) managing editor, [makes] circumstances…different. Referees may be chosen with few grounds for assessing their competence as reviewers and with no mechanism for monitoring their performance…Beyond issues of competence are considerations of reviewers' engagement with a manuscript and their affective and aesthetic responses, as noted in Silverman's [1988] sensitive analysis of peer judgment. As Silverman says, reviewers do not simply address a manuscript rationally; they respond from 'experiences, biases, knowledge, expectations, interests, and hopes' [1988, p. 364], which are, in turn, a basis for judgments rendered. Such conditions of peer review as well as quality of manuscripts received and the maturity of manuscripts upon submission vary for journals…Where circumstances of slack selection and monitoring of reviewers do exist, however, some question the very meaning of 'peer' and 'review' (Fox, 1994, 300; see also Hirschauer, 2010).

The first statement captures the normative basis on which scholars (and all others) believe peer review operates or ought to operate (Chase, 1970; Shadish, 1989). The second calls attention to the behavioral basis on which scholars actually act as participants in peer review processes. These behaviors are often invisible, disavowed, and/or discounted by others (and by participants themselves) in order to socially uphold the integrity of peer review and, by extension, the legitimacy of science.

The theme of ideal versus reality is critical since it wraps around the heart of what peer review seeks to accomplish. Paradoxically, as the process of academic peer review has matured, the problems and controversies identified in its practice have increased and broadened. We turn in the following section to how peer review is itself socially organized, which sets the stage, in the chapter's final section, to the areas of research that have challenged the ideals, and exposed many of the realities, in the contemporary practices of peer review.

The Social Organization of Peer Review

Peer review has both a social organizational role in the academic profession, as discussed above, and its own social organization, the focus of the discussion that follows. That is, in saying peer review is itself socially organized we mean that it is structured in specifiable, though often hidden, ways which results in the transparent production of publication. We consider three sets of concerns: the expansion and differentiation of peer review, the study of peer review for its labor process, and a conceptualization of peer review as an instance of "invisible interaction." Each set of concerns, in its distinctive ways, convey how peer review is socially arranged, and are reflective of the relevant literature. Further, as the discussion will reveal, these concerns exist as sources of strain in how peer review actually operates. The discovery of these strains will draw us to a still different literature, discussed in the final section of the chapter, on the dysfunctions that have developed in contemporary peer reviewing. To understand the problems that have emerged in modern peer review, it is important to provide a context that accounts for how peer review is socially set-up.

Expansion and Differentiation

Organizationally, peer review is situated in journals and academic publishing houses. Among the most noteworthy of changes over time consists of the expansion and differentiation of publishing outlets. Since the establishment of in-house journals at the Royal Society of London and the Académie Royale des Science of Paris in the late seventeenth century (Biagioli, 2002) the number of academic journals has steadily risen. In a classic study, Price estimated a doubling every 15 years in the number of journals between 1665 and 2000 (Price, 1975). Analysis of *Ulrich's Periodical Directory*, which has monitored all types of periodicals since 1932, places the number of peer-reviewed journals at 16,925 in 2002 and 23,973 in 2008 (Tenopir & King, 2009). Powell (1985) documented a similar pattern of expansion in book publication, where the number of publication companies doubled from 804 in 1954 to 1650 in 1977. Expansion is important to peer review because it signifies the magnitude of peer- reviewing activity taking place and the labor (of many types) necessary to its operation.

Expansion of the number of journals *within* fields feeds an additional and highly consequential aspect of social-organizational change: differentiation. Differentiation is manifest in two ways. First, the caliber of journals and book presses has broadened. Second, coercive pressure has intensified wherein editors compel authors to cite work from the journal in which authors seek to publish, thereby attempting to distinguish one outlet from another.

On the first point— journal and book publisher prestige— publication venues are "segmented;" that is, scholars in fields draw distinctions between general and specialized outlets on the one hand, and finer-grained distinctions in outlet quality on

the other. In all fields, publication venues are symbolically ordered by their visibility, which shapes scholars' outlet preferences. Garfield's (1955) development of journal impact factor measurement—which ranks journals according to the number of citations they receive relative to articles published in a given time period—has become widely popular in academe (Power, 1997; Tuchman, 2009). The use of "journal impact factors" likely reifies author preferences for publication outlet. This is likely so even in light of the fundamental observation that, independent of outlet, evaluating scientific quality is difficult and problematic (Seglen, 1997). Thus, in the modern era, there is little doubt that abundant good science and scholarship is published outside of the most prestigious venues, and this pattern has likely intensified with the expansion of publishing outlets (cf. Cole & Cole, 1973). The research role and institutional expectations for publication productivity have intensified over time (Hermanowicz, 2016b), but expansion of publication outlets has occurred primarily in journals considered outside of the top tier. Consequently, opportunities for academic publishing have expanded, but not among the most prestigious journals.

On the second point—coercive pressure—segmentation induces editors to manage their journals to out-compete others. Coercive self-citation refers to pressure that editors exert on authors to add citations from the editor's journal, which can be viewed as a violation of the norm of disinterestedness (Merton, 1973a), given that the communication of science is biased by the interests of an editor. Through such practices, editors "game the system." They use the peer review process to attempt to enhance their journal's prestige by improving its impact factor. In a survey of 6672 researchers in economics, sociology, psychology, and business, Wilhite and Fong (2012) found that 42 % of their respondents had experienced coercive-self citation pressure from editors during peer review; 86 % of the respondents viewed the practice as inappropriate. What is more, editors are more likely to coerce assistant and associate professors (Wilhite & Fong, 2012).

More broadly, the differentiation of journals has resulted in new forms of peer review. This has also been fostered by other forces, including long review times and delays in publication (Björk & Solomon, 2013). The journal *Nature*, for example, conducted an experiment in which papers that survived an initial editorial assessment were subjected to open peer review online in addition to standard peer view. But because of a reluctance of researchers to offer open comments, the journal deemed the experiment a failure (Greaves et al., 2006). An alternative approach to peer review, now institutionalized across a number of fields, follows the model of the journal *PLOS ONE*. *PLOS ONE*, and journals that have mimicked its format, are open-access. Its peer review process is limited to the technical validity of results and explicitly excludes judgments about the scholarly contribution of manuscripts (Björk et al. 2010). Assessment of a manuscript's scholarly contribution in this model is determined solely by whether the scholarly community ignores or cites the research.

The expansion of outlets and corresponding intensification of peer reviewing activity prompts significant organizational consequences. One consists of strain on the operation of peer review. A conservative estimate of the number of peer reviewed articles published annually is one million (Björk, Roos, & Lauri, 2009). In most

cases, successful submissions involve two or more rounds of reviews conducted by a minimum of one, but often more, reviewers before acceptance. But the number of articles published in a given year is dwarfed by the number rejected. And yet all submissions—those ultimately published and those rejected in a given cycle—are subjected to review. As Kravitz and Baker (2011, p. 1) explain: "each submission of a rejected manuscript requires the entire machinery of peer review to creak to life anew." Consequent strain in the system may be more likely in low-consensus fields, such as in the humanities, relative to high consensus fields, such as in the physical sciences, because rejection rates are higher in areas of work where scholars disagree as to what constitutes credible scholarship (Zuckerman & Merton, 1971, p. 77). Nevertheless, publication delays have increased substantially over the past several decades due in part to an increase in the volume of manuscripts to review and the time needed by reviewers to undertake the uptick in review requests (Björk & Solomon, 2013).

But submission rates and review delays themselves are found to vary by the segments of publishing outlets. Institutional expectations exact a pressure on authors to submit their work to the most prestigious journals. Thus the volume of submissions at highly-ranked journals has risen markedly. This outcome generates system inefficiencies (Sugimoto, Larivière, Ni, & Cronin, 2013). In the words of Bruce Alperts, former editor-in-chief of *Science*, bureaucratic obsession with impact factors "wastes the time of scientists by overloading highly cited journals such as *Science* with inappropriate submissions from researchers who are desperate to gain points from their evaluators" (Alperts, 2013, p. 787). Many manuscripts would be better sent initially to outlets other than the most prestigious, which by turn would help yield greater efficiencies in the operation of peer review.

Labor Process

An organizational perspective is useful for mapping the expanse of the peer review system, as well as for discerning the system's micro-level characteristics. Here we focus on the labor process underlying peer review, a view that brings into relief occupational and micro-organizational features of evaluation in scholarship and science.

For the individual author, peer review is omnipresent in scholarship and science because one's work continues to be evaluated even after it is published. When delimited to the publication process, peer review can be understood as having informal and formal stages of evaluation. During the production and completion of a scholarly manuscript, some authors ask colleagues to offer feedback prior to revision and submission to a journal; conferences provide another venue for informal peer review. Formal peer review begins with submission to a journal or press. It consists of a transaction involving "suppliers" (authors), "buyers" (journals or academic presses), and "goods" (manuscripts). Authors seek to increase their symbolic capital (Bourdieu, 1984) by publishing in journals with the highest impact, whereas editors

seek to publish papers that increase or maintain the prestige of their journals (Sugimoto et al., 2013). From one perspective, reviewers accumulate symbolic capital (noted on one's vita and journals' annual reviewer acknowledgements) when asked by editors to evaluate a manuscript: the request symbolizes recognition of specialized expertise.[2] Reviewers have less to gain in this exchange beyond exposure to new research. Accordingly, reviewers often spend little time evaluating scholarship and offering feedback. An analysis of time invested by reviewers, drawing on a survey of 276 reviewers in public health, found that reviewers spent an average of 2.4 h per paper (Yankauer, 1990). Low incentives and limited time invested in peer review raise questions about its reliability, a topic covered later in the chapter.

The labor process underlying this exchange illustrates a craft system of production characterized by autonomy, personal discretion, and informal control (Powell, 1985). Although the sequence of peer review processes can vary within and across fields and journals, formal peer review typically begins with a screening process in which an editor decides whether a manuscript is suitable for external assessment. A study of 100 randomly selected journals in 43 fields found that 10 % of journal submissions receive a "desk reject" by editors (Juhasz, Calvert, Jackson, Kronick, & Shipman, 1975); that is, the manuscript is rejected by the editor without sending it out for evaluation by external assessors.[3] In another study that included 1008 manuscripts submitted to three elite medical journals, 76.5 % of initially submitted manuscripts were desk-rejected (Siler, Lee, & Bero, 2015).

If a manuscript is selected for peer review, editors and (at times) editorial board members select external reviewers to evaluate the submission. The labor process of peer reviewing creates "small world" social networks that are decentralized, informal, fleeting, and which include a "horizon of observability". "Horizon of observability" refers to distance in a social network beyond which persons are unaware of others' roles (Friedkin, 1983). In peer review, the horizon of observability encompasses the idea that individuals are tied to one another through the peer review process with no knowledge of one's role as author or referee. What makes this particularly interesting is that reviewers and authors may be tied together through other means, such as doctoral affiliation or mutual participation on a conference panel, yet are unaware of their connection through peer review. An

[2] Requests to review, however, are not universally a sign of recognition. The role of recognition in this respect is historically contingent, pre-dating the rise of electronic submissions. Authors are now often required to indicate areas of interest when submitting a manuscript to a journal, selections that provide editors and editorial staff with a pool of reviewers to approach in the future. Thus, even a relatively unknown researcher who might have rarely published can be approached to review scholarship. Therefore, reputation and the exigencies of managing a journal likely both come into play in reviewer selection.

[3] Juhasz et al. (1975) do not specify which fields comprise their sample. They note in a comment about fields that "the conclusions of our study do not relate to any differences between the acceptance-rejection ratios between the humanistic literature and the scientific and technical literature" (p. 184).

account told by former *American Economic Review* editor, Preston McAfee, captures the point:

> At a conference, I overheard one author tell another economist that an idiotic referee reviewed his *AER* submission. He detailed all the stupid things that the referee had said, and the economist listening to the story commiserated and wholeheartedly agreed with the author, even though the commiserator was the referee in question. This referee had written a thoughtful and serious report on the paper of a friend, but the author did not appreciate the insights in the report (McAfee, 2014, p. 60).

In a survey of 94 academic psychiatrists who had submitted a paper to the journal *Psychological Medicine*, Wessely, Brugha, Cowen, Smith, and Paykel (1996) found that, of the 252 reviewers associated with the respondents' manuscripts, only 15 % were accurately identified by authors. Nevertheless, it can be the case that the research, instruments, and resources described in manuscripts belie blind review (Knorr-Cetina, 1981). Crane (1967), for example, discussed the presence of "invisible colleges" wherein the identity of authors is recognizable by their style, use of data, and theoretical orientation.

Further, the proliferation of publication has rendered calls to review more challenging (Hermanowicz, 2016a, 2016b). The editor of the *American Journal of Sociology* has explained that 10 people are customarily asked in order to yield three reviews, but in cases the requests have climbed to as high as 17 (Abbott, 2011). This was not the case prior to using electronic means to solicit reviewers. Receiving a packet of a manuscript and forms in the mail, reviewers evidently felt more compelled to do the review. Receiving an email request, reviewers are more readily able to push a button and decline (Abbott, 2011).

Calls to review are more frequent, faculty workloads have increased (Jacobs & Winslow, 2004), and reviewing in journals is unremunerated (except honorifically) (Kravitz & Baker, 2011). In principle, editors select qualified reviewers who have expertise in some aspect of the scholarship under consideration, but research on reviewer selection exposes the influence of other factors that shape this work. Social proximity (Hamermesh, 1994), status (Stossel, 1985), and career stage (Fyfe, 1994; Guest, 1994) shape the selection of reviewers and potentially the quality of the review. Stossel (1985) found that high-status reviewers are disproportionately more likely to refuse to review manuscripts. This means that editors often must rely on younger and less experienced colleagues to evaluate submissions. Yet, eminence and seniority do not assure high-quality reviews. Editorial commentaries on peer review practices reveal complaints about the quality of reviews submitted by high-status and senior faculty (Finke, 1990; Judson, 1994). Quantitative analyses that use length and elaborateness of a review as a *proxy* for review quality suggest that the highest proportion of high quality reviews are, indeed, provided by younger and lower-status colleagues (Stossel, 1985). And while it is reasonable to posit that reviewer selection is shaped by subfield, given that larger subfields offer larger pools of potential reviewers (e.g. there are more particle physicists than string theorists and more sociologists of stratification than historical sociologists), disciplinary and subfield differences in reviewer selection remain unprobed territory in research on peer review.

Invisible Interaction

Publication results from an iterative process that includes exchanges between authors and reviewers (Teplitskiy, 2015). While gratitude is paid occasionally to "anonymous reviewers" in acknowledgements of published manuscripts, tangible publications occlude the interaction between authors and reviewers, much like other cultural products (such as musical instruments or pieces of art) tend to conceal the work of others who help produce them (Becker, 1984). Hood (1985) has referred to this social pattern as the "lone scholar myth." Knorr-Cetina (1981) has described how scientific papers entail a "hidden monodrama" in which authors argue with specific others.

Existing work characterizes the relationship between authors and reviewers in two ways. Some scholars depict a *collaborative* relationship between authors and reviewers in which reviewers and editors attempt to "coach" authors by providing feedback on how to deal with flaws in theory, methodological issues, and implications of findings (Cummings, Frost, & Vakil, 1985; Jauch & Wall, 1989). Other scholars depict an *oppositional* relationship in which reviewers and editors challenge authors to meet expectations of scholarly merit (Crane, 1967; Strang & Siler, 2015). Reviewer characteristics are rarely studied comparatively, but one study shows that philosophers' reviewers are more negative and likely to recommend rejection than psychologists' reviews (Lee & Schunn, 2011), suggesting "toughness" may vary by field. Describing the book publication process, Coser, Kadushin and Powell (1982:198–199) note that conflict is endemic to the publishing process. In his study of two commercial publishers of scholarly books, for example, Powell (1985, p. 64) described how authors get "caught in the crossfire" of battles between different departments of the publishing house. The review process at journals can operate similarly.

Two characteristics of the interaction between authors and reviewers are noteworthy. First, a power/dependence differential exists in which authors have something to lose, while reviewers have little to gain (Strang & Siler, 2015). Second, authors, reviewers, and editors may share a commitment to the production of knowledge, yet each party brings varied sets of interests to the exchange. Reviewers' interests may potentially shape how they evaluate an article. For example, reviewers can promote their own reputation by pressuring authors to cite their work (Wilhite & Fong, 2012). They may try to protect their reputation within an area by rejecting findings that contradict or challenge their research (Abramowitz, Gomes, & Abramowitz, 1975) or by delaying the publication of research that competes with their own (Grouse, 1981). In the most egregious circumstances, scientists have expressed concern that peer reviewers have stolen their ideas and published them elsewhere (De Vries, Anderson, & Martinson, 2006).

Knorr-Cetina's (1981) ethnography of a major research center integrates collaborative and oppositional views of author-reviewer exchange. Her work sought to demonstrate that publications are the result of a negotiated compromise. The publication of scholarship necessarily entails writing against others and, while reviewers

may be friends or acquaintances with an author in an overlapping network, they are also competitors who work on similar topics with stakes of their own (Knorr-Cetina, 1981).

The negotiations between authors, editors, and reviewers have been investigated only sporadically. For example, Simon, Bakanic, and McPhail (1986) examined complaints to editors from authors using 250 manuscripts submitted to the *American Sociological Review*. They concluded that complaints are rare, most frequently focus on perceived reviewer incompetence, and rarely persuade editors to reconsider submissions.

In other work, researchers have shown that reviewers and authors in the social sciences negotiate the meaning of a paper's findings, rather than the evidence. Using content analysis of published submissions and a survey of authors in *Administrative Science Quarterly* between 2005 and 2009, Strang and Siler (2015) introduced the idea of "revising as (re)framing." They found that theory sections of papers were intensively reworked and hypotheses were altered, but changes to measures and analyses were limited. Teplitskiy's (2015) comparison of conference papers later published in the *American Sociological Review* and *Social Forces* similarly showed that manuscripts change more in their theoretical framing than in their data analyses. How the notion of "revising as (re)framing" applies to other fields remains empirically unexamined. One might hypothesize that theoretical reframing is most operative in low- and medium-consensus fields where scholars are more likely to disagree about appropriate hypotheses and interpretations, and least operative in fields characterized by high-consensus, given the high levels of agreement about theoretic frameworks in such fields.

An examination of peer review for its own social organization has drawn attention to the expansion and differentiation of publication outlets that have directly affected peer review, to the study of peer review for its labor process, and to an understanding of peer review as instantiations of "invisible interaction." While publication may enshrine works of scholarship with a sense of objectivity, these three sets of concerns underscore the point that peer review is a socially susceptible and socially sustained series of processes. To these ends, publications that result from "successful" peer review are negotiated products that carry the interests of authors, reviewers, and editors. Other works are, of course, rejected as a result of the "same work" of interests. They are shelved, and then some of them are sent elsewhere, and a subset of these are eventually published or rejected again, repeating the processes until authors, reviewers, and editors at some point, at some outlet, are of at least some general accord—to finally publish, or to finally give up in finding a home.[4]

[4]The length of time that authors take to re-submit manuscripts for consideration elsewhere is a question open to empirical inquiry. Length of time may function, for example, from reviews received and decisions to revise before sending manuscripts to different outlets. Length of time may also function by career stage; the younger the stage, the lesser time on the shelf. In the earliest stages, shelves may be altogether short: anecdotal evidence suggests that many junior scholars are advised always to keep completed manuscripts under review, under the premise that they will "hit" somewhere. The premise itself underscores the socially situated variability of peer review processes (i.e., manuscripts of all sorts will be accepted somewhere).

While one view holds that these processes are functional and act in service to a larger system of science, this is but one view. By turn, these processes are stamped with the imprimatur of human behavior. Thus a circumspect consideration of peer review must also take into account social-behavioral realities in the practices of peer review. These considerations turn our attention to dysfunctions that arise in how modern peer review operates.

Dysfunctions of Peer Review

Given that peer review processes are organized socially, in which there may often exist competing interests and conflicts, strongly contrarian attitudes about peer review have developed, not only in the informal life and experience of academe but also in a critical literature that has burgeoned over the last three decades. (This period coincides with an ascendance in the usage and in the consequences of peer review; publication outlets, calls to review, and expectations to publish, have all grown.) The ascendance of contrarian views has occurred even in light of the long line of points established at the outset that accounted for how and why peer review came to constitute nothing less than "the linchpin of science." Thus one reads, in but one of many editorial columns (this coming from the acclaimed scientific journal *Nature*), that peer review is "the least imperfect way of upholding the quality of scientific publications" (Dewitt & Turner, 2001, p. 93). Such a statement in 1950, let alone an industry of such sentiment, would be unimaginable. A considerable body of research has examined operational machinations of peer review in publication. Through a variety of lenses, much of this research confronts the question of whether peer review is a reliable arbiter of knowledge. This section of the chapter examines the operation of peer review by considering research on its reliability and impartiality. Studies reveal a general agreement among researchers that reviewers themselves rarely agree about the quality of manuscripts. The remaining bulk of the section addresses factors identified in research that seek to explain why the reliability of peer review appears to be so low.

Reliability

"Two scientists acting unknown to each other as referees for the publication of one paper usually agree about its approximate value," asserted Polanyi (1946, p. 51). Polanyi's claim, if validated empirically, contends that peer review is a highly reliable process of evaluation. The crux of reliability rests on a simple premise: peer review cannot be a valid means of assessing scholarly or scientific quality if independent referees do not reach similar conclusions about the merit of manuscripts. The earliest research on reliability implied that peer review is, in fact, conducive to reviewer agreement in the assessment of manuscripts. Analyzing 172 manuscripts

submitted to *The Physical Review* between 1948 and 1956, Zuckerman and Merton (1971) found that reviewers recommended the same decision in all but five cases; only minor differences emerged in comments about revisions necessary for publication. Research examining biomedical science (Orr & Kassab, 1965) and sociology (Smigel & Ross, 1970) produced similarly positive depictions of reviewer agreement.

Serious questions about the reliability of peer review, however, were subsequently raised by quasi-experimental studies that identified major failures of the peer review process. Rather than testing reliability by describing frequency of agreement among reviewers of published papers (Orr & Kassab, 1965; Smigel & Ross, 1970; Zuckerman & Merton, 1971), studies in this vein audit the review process by testing how well reviewers detect blatant errors. One such failure includes bias (Mahoney, 1977; Peters and Ceci 1982), a topic addressed shortly. Another such failure involves an inability to recognize fallacious submissions (Sokal, 1996). The "Sokal hoax," for example, involved New York University physicist Alan Sokal, who devised a manuscript comprised of unintelligible and nonsensical writing—presented as a postmodern argument about math and physics. He successfully published the paper in the journal *Social Texts* (Sokal, 1996). While the goal of the paper was to critique postmodernism, the episode revealed how editorial review failed to detect a completely erroneous and illogical contribution.

Similarly, in collaboration with molecular biologists at Harvard, Bohannon (2013) devised a paper on cancer cells that was written in a credible manner but which included obvious fatal errors. The manuscript was submitted under a series of fictitious author names affiliated with fictitious African institutions to 304 open-access journals, of which only 98 rejected the paper. Of the 106 journals that provided evidence of peer review, 70 % accepted the paper (Bohannon, 2013). The Sokal (1996) and Bohannon (2013) ploys speak most explicitly about *editorial* peer review and, in the case of Bohannon (2013), open-access publishing. Similar events have occurred in peer review of conference papers (McLachlan, 2010). While some of these events attract more popular attention than others, their message is meant as general: the peer review process can be a weak filter for flawed, fabricated, and even non-sensical work. Such events point to significant shortcomings of scholarly evaluation and the limits of peer review as a method for detecting fraudulent work (Fox, 1994).

Quantitative analysis of reviewer agreement, or inter-rater reliability, is the primary means by which researchers have assessed the reliability of peer review. Statistical approaches to examining inter-rater reliability vary across studies. One of the main techniques used is intraclass correlation, where a correlation of 1.0 indicates full agreement between reviewers. Most studies of this kind identify low levels of consensus in reviewer recommendations. In studies with relatively large samples of manuscripts, researchers have found intraclass correlations of .12 in management (Starbuck, 2003), .23 in law (Cicchetti, 1991), .29 in sociology (Hargens & Herting, 1990), and in psychology—one of the most active areas of research on inter-rater reliability—estimates vary from as low as .08 (Cicchetti & Eron, 1979) to .34 (Cicchetti, 1991). A meta-analysis of 48 studies of inter-reviewer reliability similarly concluded

that reviewer agreement tends to be low and that studies reporting high levels of agreement tend to be based on small sample sizes (Bornmann et al., 2010).

These measures, however, reflect *overall* agreement among reviewers, meaning they are averaged across all possible recommendations including "accept," "reject," and "resubmit." As Cicchetti's (1991) review of peer review in psychology, sociology, medicine, economics, and the physical sciences demonstrates, reviewer consensus is more nuanced than overall averages depict. In decisions about manuscripts to accept, the level of agreement among reviewers varied between 44 and 66 %, whereas agreement about manuscripts to reject varied between 70 and 78 % (Cicchetti, 1991). Still, even accounting for the type of editorial recommendation, reviewer agreement tends to be low in general across many fields, and agreement about acceptance is lower than agreement about rejection (Campanario, 1998; Lee et al., 2013).

It is important to note that these studies fail to control for key factors, such as the experience of the reviewer, the experience of the editor, and the technical or theoretical complexity of a submission. Extant studies of reliability have also generally focused on a highly selective subset of journals; the above studies examined prestigious journals. Although this approach is strategic, insofar as it seeks to capture peer review in its most rigorous form, the norms of review at such journals may reflect a tendency to emphasize the finding of faults in submissions. Reviewing behavior at prestigious journals is customarily circumscribed by the norm: "when in doubt, reject" (Zuckerman & Merton, 1971). By contrast, different evaluation norms may operate at journals with high acceptance rates, conveyed in the converse aphorism: "when in doubt, accept" (Zuckerman & Merton, 1971).

While reviewer agreement tends to be low, one might expect variation in reliability across fields due to different degrees of consensus (Braxton & Hargens, 1996). As Hargens (1988, p. 147) has explained: "When scholars do not share conceptions of appropriate research problems, theoretical approaches, or research techniques, they tend to view each other's work as deficient and unworthy of publication." Sociologists, for instance, lack precise definitions of concepts and consequently rely on multiple measures of key ideas, suggesting that reviewer disagreement would be more prevalent in sociology than, for example, in chemistry, a higher consensus field. Cole, Cole and Simon (1981) found reviewer consensus roughly equal in proposals for funding in economics, biochemistry, and solid-state physics—but all are high-consensus fields. In their meta-analysis, Bornmann et al. (2010) found no relationship between discipline and reviewer agreement, but their categories—such as "natural sciences," "economics/law," and "social sciences"—may have been too broad to adequately assess this relationship. An absence of studies that assess differences in reviewer agreement between variegated fields limits our ability to fully sort out a relationship between field consensus on the one hand and reviewer reliability on the other.

Bias

Despite expansive lists of normative criteria that scholarly communities believe *should be used* in peer review—criteria that emphasize issues such as the relevance of literature review and the adequacy of research methodology (Cicchetti, 1991; Ramos-Alvarez, Moreno-Fernandez, Valdes-Conroy, & Catena, 2008)—little is actually known about how reviewers conceive of and evaluate scientific or scholarly merit (for an exception, see Guetzkow, Lamont, & Mallard, 2004). But a wealth of research has, however, attended to the influence of factors *unrelated* to merit. While redundant studies of reliability populate the literature, the topic of bias has assumed the greatest share of researchers' attention to the peer review process. It is, then, no coincidence that *bias*—in its various manifestations—consumes a proportional share of the present review.

Conceptually, whether or not reviewers are impartial evokes a norm of "universalism," an institutional imperative, which holds that scholarly and scientific contributions should be evaluated according to "preestablished impersonal criteria" and without consideration of social attributes of authors, such as age, gender, race, nationality, past or present institutional affiliation, or other subjective characteristics (Merton, 1973a, pp. 270–271). Conformity to the norm of universalism is understood as a condition of scientific debate within academic communities, unbiased data interpretation, and intellectual freedom.

Universalism is functionally understood to serve the goals of science and scholarship: to extend knowledge free from bias, for it is only then that a legitimate furtherance of knowledge can occur. The impartiality of peer review is an important feature of higher education's distinctive societal mandate: lack of impartiality entails an erosion of academe's mandate and its status (Merton, 1973a), centrally because legitimate knowledge can only be promulgated in the absence of irrelevant interests. Impartiality may also perpetuate orthodoxy in the cognitive structure of science, as when reviewers reject research that is inconsistent with a theoretic tradition to which they adhere (Travis & Collins, 1991). What is more, when perceptions of bias in peer review are endemic, and these conditions may vary from field to field for reasons associated with consensus as suggested above, cynicism about career goals, about institutions, about fields, about colleagues—indeed much of academic life—can capture and contaminate a great share of a scholar's imagination (Gillespie et al. 1985; Hermanowicz, 2009). Such a dynamic is, of course, contrary to the goals of academe.

Contemporary research on the operation of peer review originates from an agenda established primarily by sociologists in the "Columbia school" of the sociology of science, consisting of Robert Merton and his students (Calhoun, 2010; Hess, 1997, especially pp. 52–80). Work in this tradition focused on the question: "Is science universalistic?" In doing so, it sought extensive evaluation of the role of particularism. Overall, this early wave of research found that peer evaluation generally operated universalistically by demonstrating that reviewers make decisions based on scientific quality rather than on particularistic characteristics such as institutional

affiliation (Cole et al., 1981; Cole, Rubin, & Cole, 1978; Crane, 1967; Zuckerman & Merton, 1971). Nevertheless, even within this tradition of work, there have occurred highly significant changes in view about universalism and particularism in the creation of scientific knowledge (e.g., Cole, 1992; Long & Fox, 1995).

More recent research on peer evaluation has problematized the notion of scientific quality by arguing that definitions of quality are not unequivocal (cf. Guetzkow et al., 2004). The contemporary literature comes full circle from its origins: it examines sources of bias that undermine decisions based on intellectual merit. Whereas once upon a time universalism was virtually presupposed, now it is cast as "naïve realism" (Cole, 1992). The occurrence of particularism, the usage of irrelevant factors to assess contributions, carry the day. It is to this large literature that we now turn. We divide the discussion into the two broad forms of bias that the literature addresses: cognitive particularism and social particularism.

Cognitive Particularism

Travis and Collins (1991) developed the term *cognitive particularism* to describe how peer reviewers make decisions based on their adherence to specific schools of thought. This form of particularism is aptly captured in Francis Bacon's treatise on the interpretation of nature, *Novum Organum*:

> It is the peculiar and perpetual error of the human intellect to be more moved and excited by affirmatives than by negatives; whereas it ought properly to hold itself indifferently disposed toward both alike (Bacon, 1620, p. 20).

Cognitive particularism is characterized by a rigidity that biases reviewers against epistemological approaches inconsistent with their own. Three manifestations of cognitive particularism are found in the research literature: confirmation bias, positive outcome bias, and conservatism.

1. Confirmation Bias Confirmation bias refers to reviewer bias against manuscripts that include results inconsistent with a reviewer's theoretic perspective. In a classic deception study, Mahoney (1977) asked 75 reviewers to assess the same fictitious psychology experiment. Mahoney sent reviewers an identical introduction, methodology, and bibliography, but systematically altered data presentation and interpretation such that reviewers evaluated papers that generated either positive, negative, mixed, or no results. Referees were selected from a list of guest reviewers from a psychology journal associated with a specific intellectual perspective (applied behaviorist psychology). Judgments about the quality of the paper were higher when the results conformed to the theoretic perspective of reviewers and lower when results were incongruent with reviewers' theoretic perspective (Mahoney, 1977).

In a study of peer review in social work and allied fields, Epstein (1990) sent a contrived research paper to 110 journals and found that journals were more likely to accept papers that presented findings consistent with the mandate of social work, as

opposed to papers that reported findings perceived as contrary to the field's mandate. Similar patterns of confirmation bias have been documented in medicine (Ernst, Resch, & Uher, 1992).

2. Positive Outcome Bias Another manifestation of cognitive particularism, related to, but analytically distinct from confirmation bias, consists of positive outcome bias, sometimes simply called "publication bias" (Lee et al., 2013; Olson et al., 2002). In this occurrence, editors and reviewers only recommend papers for publication that report statistically significant results (independent of whether results are congruent with reviewers' theoretic perspective). By one view, bias toward positive results is beneficial to science. In *The Art of Scientific Investigation*, Sir William Beveridge (1950, p. 25) argued that inability to demonstrate a hypothesis does not prove that one's presupposition is incorrect, and consequently, "it is a commendable custom usually not to publish investigations which merely fail to substantiate the hypothesis they were designed to test." Beveridge (1950) lauded the rather drastic practice of destroying records of "negative experiments." Others who have come to work in a more mature period of science have exercised concern over this "emptying of the file drawer" (Rosenthal, 1979). A purging of work that failed to yield positive outcomes leads, in its own way, to a lack of awareness within scholarly communities of important non-findings.

More to the point, the advancement of knowledge is predicated on a self-correcting process (Knight, 2003); nonsignificant results that do not conform to expectations are of considerable importance. Scholars have drawn concern around three consequences of bias against nonsignificant results. First, misallocated effort, time, and money can result when researchers pursue lines of inquiry unaware of prior efforts in their disciplinary communities that resulted in unpublished nonsignificant results (Campanario, 1998a).

Second, researchers may abandon serious theory-building by anticipating that reviewers will reject studies reporting nonsignificant results. Instead, they preoccupy themselves with statistical significance. As Dar (1987, p. 149) explains:

> When passing null hypothesis tests becomes the criterion…for journal publications, there is no pressure on the… researcher to build a solid, accurate theory; all he or she is required to do…is produce statistically significant results.

Third, researchers have expressed concern that bias toward positive results may lead to research misconduct. In his research on the abundance of positive results in medical studies, Weisse (1986, p. 23) has noted that he was "struck by the predominance of investigators with positive findings, with the naysayers in the distinct minority…It had seemed, at times, that the only way to get ahead is to be a perpetual yes-man." Hersen and Miller (1992) assert that bias towards positive results may encourage fabrication or falsification of data because of pressures to publish, and positive outcomes are perceived as standing the greatest chance of publication. This behavioral pattern is consistent with the sociologically-based strain theory of misconduct, which posits that researchers unable to achieve success through legitimate

means will turn to illicit means to advance themselves professionally (Zuckerman 1988).

More general evidence has arisen to lend credence to positive outcome biases (Emerson et al., 2010). In a survey of 429 reviewers at 19 leading journals in management and the social sciences, Kerr, Tolliver, and Petree (1977) found that 28 % of reviewers indicated they would likely reject an article if its results did not approach statistical significance. Sterling's (1959) foundational analysis of publications in four psychology journals revealed that 97 % of the articles published during 1 year reported positive results. Modern replications similarly demonstrate that an exceedingly small proportion of articles in psychology report nonsignificant results (Coursol & Wagner, 1986; Greenwald, 1975; Shadish, Doherty, & Montgomery, 1989; Smart, 1964). Weisse's (1986) analysis of 408 articles in three medical journals found that, depending on the journal, not less than 80, and as much as 90 % of published articles reported positive results.

Change over time further indicates support of positive outcome biases. Fanelli's (2010) analysis of 4600 papers published across a spectrum of fields between 1990 and 2007 reached noteworthy conclusions. The frequency of papers that report positive findings increased from 70 % in 1990 to 86 % in 2007, roughly a 6 % increase each year. The rate of increase differed by field; social science outlets registered the greatest increase relative to the physical sciences, a pattern that is likely consistent with points about field consensus raised previously (Braxton & Hargens, 1996). It may be easier to "generate and sell" positive results in fields whose definitions of what is "positive" are more fluid.

3. Conservatism Finally, "conservatism" constitutes a means by which cognitive particularism is expressed in processes of peer review. In contrast to confirmation bias (which underscores reviewers' theoretic stances) and positive outcome bias (which underscores reviewers' devaluation of nonsignificant statistical results), conservatism arises as a bias when reviewers reject otherwise worthy submissions that are judged to run counter to convention. Here we follow the classification schemes in the peer review literature (Lee et al., 2013; Shatz, 2004) in distinguishing between these three modes of cognitive particularism. Scholars' use of the term confirmation bias focuses primarily on theoretic positions, while the term conservatism is applied more generally to innovation, change, and risky perspectives.

On the one hand, conservatism may be considered allied with peer review: "we fully expect our theories to encounter objection along the way that we cannot easily answer, and we are not expected to crumble in the face of a problem" (Shatz, 2004, p. 84). In this vein, conservatism is consistent with the so-called "norm of organized skepticism." Merton described this norm as a "system of institutionalized vigilance…[in which] scientists are at the ready to pick apart and appraise each new claim to knowledge" (Merton, 1973a, p. 339). But to speak of conservative *bias* is to take normative behavior to unwarranted excesses.

Thus the fault refers to "bias against groundbreaking and innovative research" (Lee et al., 2013, p. 9). Chubin and Hackett (1990) similarly argue that reviewer tolerance for innovativeness is limited. All the points are consistent with Kuhn's

groundbreaking treatise on normal science; it often takes a revolution to change a paradigm (Kuhn, 1962).

In this respect, conservative bias may be apt to arise in scholarship perceived as risky, unorthodox, and when new scholarship would seem to upset well-established paradigms of thought. While conservatism can arise at a variety of instances in peer review processes, it may be especially apparent in the arena of proposals for research funding since, by definition, purportedly new ideas, theories, methodologies, and/or techniques are placed under review.

Many more scholars have written to attest to an objection about conservative bias than to empirically document its occurrence. Empirical evidence is found primarily in examples of important scientific papers (indicated by prizes and citations) that were rejected one or more times but went on to receive wide recognition. Writing about Hannes Alvén, whose research on magnetohydrodynamics resulted in the 1970 Nobel Prize in physics, Dressler (1970, p. 604) noted that Alvén's "ideas were dismissed or treated with condescension; he was often forced to publish his papers in obscure journals, and he was continually disputed by the most renowned senior scientists working in the field of space physics." It is not difficult to locate similar examples of rejected papers that later became award winning classics (Leahey & Cain, 2013; Zuckerman, 1977). Campanario's (1993) novel study of "citation classics" culls the many (but by no means exhaustive list of) papers that were "too innovative for their time" to realize ready acceptance and publication.

Social Particularism

Social particularism encompasses bias that stems from the use by assessors of functionally irrelevant attributes of authors. These attributes run a social gamut: institution, race, gender, political party affiliation, religion, sexual orientation, gender identity, age, and so on. This form of particularism rests on assumptions that members of a particular group conduct research that is inferior or superior to members of other groups. Concern about this form of bias is pervasive in two senses: as a topic of scholarly inquiry within the purview of peer review, and as a subject of personal experience. With regard to the first point, its validity shall be self-demonstrating by the coverage of the literature that follows. With regard to the second, its validity is borne empirically. In a survey of nearly 4000 academics in the United States, Morton and Price (1986, p. 1) found that fully three out of four respondents viewed peer review as biased—particularly in favor of "established scholars"—and fully half believed that major reform of peer review was necessary. In this sense, academics' personal concerns about peer review "endorse" the operation of the "Matthew Effect," which holds that disproportionate recognition is conferred upon already-established or recognized scholars (Merton, 1973c).

Social particularism is most likely to arise in single-blind reviewing. The reason is straightforward: reviewers are exposed to the authors' identities of the manuscripts under evaluation. Given the frequent usage of single-blind reviewing in the life sciences, physical sciences, and engineering (Ware & Monkman, 2008), one

might speculate that the conditions for social particularism are heightened in these fields. Single-blind reviewing is, however, not a necessary condition for social particularism's occurrence. Conference presentations publicly signal an association between pieces of work and authors' identities; panel reviewers of research funding learn who is conducting particular projects and may end up reviewing resulting manuscripts; job applicants often disclose manuscripts under review on their vitas not knowing a reviewer is a member of the search committee; "invisible colleges," referenced earlier, enable colleagues to know about the projects and manuscripts of others; and an increasing number of authors make no effort to conceal—and in today's age even promote—unpublished manuscripts through websites, electronic mail, professional list serves, and other social media. In short, there are a number of paths by which reviewers glean author characteristics—characteristics that have nothing to do with idealized "preestablished impersonal criteria" (Merton, 1973a, p. 269) that normatively may be thought to constitute the bases of review. Reflecting currents in the literature, three key sources of social particularism in peer review are discussed in turn: institution, gender, and geography.

1. Institution Institutional bias encompasses reviewer partiality toward authors' present and/or past employing organizations and/or locations of educational training. Such bias is reflective of a stratification of colleges, universities, centers, and institutes and a corresponding social assignment of organizational status. Biases result when individual abilities—and the scholarly merit of manuscripts— are imputed by organizational status. Different types of institutions possess varied resources that, by turn, facilitate and constrain scholarly roles, including the quality and quantity of publication productivity (Hermanowicz, 1998). Reviewer judgments may be influenced on the basis of the perception of such organizational attributes.

In early work on the topic, Crane (1965) found that highly productive scientists at high status universities were more likely to receive recognition than highly productive scientists at less reputable universities. Powell's (1985) ethnography of major book publishers captures well the operation of institutional bias through the remarks of an editor-in-chief:

> If I received…a manuscript on sociological theory from someone at East Delta State University, I would not consider it for a minute. If the manuscript is any good, why isn't the author at Berkeley…? (Powell, 1985, p. 95).

An early approach to examining institutional bias involved measuring the institutional composition of published research in particular fields. An analysis of 484 articles published in the *American Economic Review* between 1950 and 1959 showed that most authors held appointments at elite universities. Authors located at the University of California—Berkeley, the Massachusetts Institute of Technology, Stanford University, and the University of Chicago accounted for one-fifth of the papers (Cleary & Edwards, 1960). Similar observations have been made by examining the doctoral origins of contributors to top journals. Cleary and Edwards (1960) found that 45 % of all contributors to the *American Economic Review* received their

doctorates from Harvard, Columbia, and Chicago. Wanderer's (1966) analysis of contributors to the *American Sociological Review* between 1955 and 1965 showed that 40 % of articles appearing during that period were written by authors with doctorates from Chicago, Columbia, Harvard, and Michigan. Goodrich (1945) analyzed manuscripts received and published at the *American Sociological Review* over a period of 16 months, and found that manuscripts submitted by authors at prominent departments of sociology were accepted in greater proportion than those submitted by authors at less prominent departments.

Crane (1967) adopted a comparative approach to infer whether reviewers' awareness of authors' institutional affiliations influenced recommendations for publication. The analysis focused on a two decade period during which the *American Sociological Review* adopted a blind review policy, allowing a comparison between single-blind reviewing and double-blind reviewing at one journal. Crane also compared these patterns to those of the *American Economic Review*, which utilized solely a single-blind system in which reviewers knew author and institutional identity. At the *American Sociological Review*, Crane (1967) found that the proportion of authors from major universities actually increased after double-blind review was adopted. When comparing data between sociology and economics, Crane (1967) found that diversity in the doctoral origins of editors—having editors whose doctorates were from less-prestigious universities, for example—had greater influence on the proportion of articles appearing by scholars at less-prestigious universities. This "editor effect," operating in the opposite direction of most patterns of institutional bias, out-weighed effects in the usage of single- verses double-blind reviewing.

Utilizing a quasi-experimental design, Peters and Ceci (1982) selected twelve articles written by faculty members at prestigious institutions. The articles had been published in the 18 to 32 months leading up to their study. Each article was published in a reputable psychology journal with high citation rates, and all of the journals used non-blind reviewing. Peters and Ceci resubmitted the twelve manuscripts to the same journals that published them. Before doing so, they altered both the authors' names (but not their gender) and status of their institutional affiliations. They made only cosmetic changes to titles, abstracts, and beginning paragraphs. Three of the papers were detected as existing publications; one was accepted for publication; eight were rejected. Given that the key difference between the original and resubmitted manuscripts was institutional status, Peters and Ceci (1982) turned to institutional bias as the explanation of the assessment differences. The work sparked animated controversy and commentary for the bias it seemed to so transparently reveal (Weller, 2001, pp. 220–221; Harnad, 1982). Others, in related work, have drawn similar conclusions about institutional bias (Shatz, 2004; Lee et al., 2013). Subsequent research has sought to investigate whether institutional reputation carries more weight in reviewers' evaluations than the actual research under review. Examining reviewers' recommendations and editorial decisions for manuscripts submitted to the *Journal of Pediatrics*, which uses single-blind reviewing, Garfunkel, Ulshen, Hamrick, and Lawson (1994) found that lower institutional rank was significantly associated with lower rates to recommend publication.

2. Gender While research on institutional bias crosses decades, scholars have relatively recently begun to investigate matters of gender bias. The increased attention may stem in part from the changing demographic composition of the academy (National Science Board, 2010). Distinctions begin at the reviewing outlets themselves. Cho et al. (2014) found that between 1985 and 2013, 16 % of the subject editors were women at ten highly regarded journals in the biosciences. In an analysis of 60 top-ranked journals in the field of medicine, Amrein, Langmann, Fahrleitner-Pammer, Pieber, and Zollner-Schwetz (2011) found that 17.5 % of all editorial members were women. While gender representations of editorial boards constitute distinctions, they are not in and of themselves indicative of bias.

Some studies have investigated gender bias by ascertaining whether the evaluation of manuscripts differed by the usage of a female or male author name. In general, these studies have found that, whether rated by males or females, male authors tend to receive more favorable assessments (Goldberg, 1968; Levenson, Burford, Bonno, & Davis, 1975; Paludi & Bauer, 1983; Ward, 1981). But, in studies employing a similar methodology, the differences have not been statistically significant (Levenson et al., 1975; Ward, 1981).

Lloyd (1990) obtained the names of researchers listed on the editorial boards of five behavioral psychology journals (a female dominated field). Lloyd sent different versions of a fabricated manuscript to 65 men and women reviewers, altering only a fictional male or female author name. Male reviewers recommended acceptance of 30 % of male-authored manuscripts and 21 % of female-authored manuscripts; in both instances the acceptance rates are low, and the differences fail to provide clear indication of bias. By contrast, female reviewers recommended acceptance of 10 % of male-authored manuscripts and 62 % of female-authored manuscripts; in one instance the acceptance rate is very low, and in the other, high enough to flag a substantial difference in rates. Female reviewers accepted significantly more female-authored than male-authored papers. Lloyd's work (1990) illuminates the possible existence of gender bias, though perhaps not in ways many might suspect. The study suggests that gender bias may operate as a function of whether a field is considered masculine, feminine, or gender neutral.

Comparisons of *acceptance and rejection rates* by gender do not reveal **strong** evidence of bias (Ceci & Williams, 2011; Weller, 2001. Many studies have typically found slightly higher acceptance rates for articles published by men compared with women, but the differences rarely achieve statistical significance. In economics, studies of authorship patterns at fourteen journals found that women—with or without male coauthors—had higher acceptance rates under blind peer reviewing than at journals using non-blind reviewing. But no *statistically significant* differences in acceptance were observed between men and women (Edwards & Ferber, 1986; Ferber & Teiman, 1980; see also Blank, 1991).

Studying 592 manuscript submissions at the social science journal *Human Organization* over a four year period, Bernard (1980, p. 369) concluded that "aggregated data indicate no gender bias for getting into print..." In one of the largest analyses of gender bias in peer review, the editors of the *Journal of the American*

Medical Association reported observable gender differences in editor and reviewer demographic representation, but none that held implications for manuscript acceptance (Gilbert, Williams, & Lundberg, 1994).

In 2001, the journal *Behavioral Ecology* changed from non-blind to double-blind reviewing. Budden et al. (2008) compared the representation of female authored publications, both before and after the shift, and in comparison to five similar journals. The authors found that the acceptance rate for female first-listed authors increased by roughly 8 % in the 4 years after the onset of blind peer review. Furthermore, the researchers found that the increase in the proportion of female authors in *Behavioral Ecology* was greater than at five other journals that used single-blind reviewing. Critiques of this study emerged on statistical grounds, namely that Budden et al. (2008) were unable to rule out the possibility that the increasing percentage of articles published by women was the result of an increased number of articles submitted, and not decreasing gender bias (Webb, O'Hara, & Freckleton, 2008; Whittaker, 2008). More broadly, analyses indicate that manuscripts by female life scientists are not rejected disproportionately, whether at single-blind journals, such as *Journal of Biogeography* (Whittaker, 2008) and *Cortex* (Valkonen & Brooks, 2011), or at double-blind journals, such as *Biological Conservation* (Primack, Ellwood, Miller-Rushing, Marrs, & Mulligan, 2009).

3. Geography An emergent area of research on social particularism in peer review examines bias associated with geography. This instance of bias is seen to operate in two primary ways. First, reviewers are thought to use an author's country as a proxy for evaluating the quality of scholarship. In the natural sciences, for example, reviewers could register bias against research from a region outside the global science infrastructure. Reviewers can make an assumption that the instruments and technologies needed to perform cutting-edge research are unavailable in a particular region, or that adequate training for scientists is unavailable.

Second, journals are thought to favor the publication of manuscripts from authors located in the same nation as the journal. Here it may also be noted that some scholars have called attention to the existence of language bias (Herrera, 1999), but little empirical work supports the claim (Loonen, Hage, & Kon, 2005).

Much of the research on geographical bias is situated in analyses of peer review in medical scholarship. One of the first studies of geographic bias investigated the hypothesis that journals favor the publication of scholarship coming from the country in which the journal resides (Ernst & Kienbacher, 1991). Ernst and Kienbacher (1991) examined articles published in 1990 at four journals in the field of physical medicine and rehabilitation; the journals were respectively based in Britain, Sweden, the United States, and Germany. They found that only the United States journal exhibited a skewed regional representation of authors in this specific scientific field. The three journals published in Europe had roughly equivalent representation of native and foreign authors. In the U.S. journal, 79 % of authors had appointments in the United States, while 21 % came from other regions. Skewed representation of authorship is not, of course, equivalent to bias. That is, we are not able assess the merits of the submissions to the U.S. journal independent of the regional locales of

authors. It is conceivable that the journal simply accepted the best work submitted to it, and the authors of the manuscripts clustered in the U.S. It is equally likely that, given national differences in science infrastructure, regions vary in the volume of research produced.

An additional perspective is offered by a retrospective study of articles submitted to the U.S. journal *Gastroenterology* in 1995 and 1996 (Link 1998). The study sought to determine whether U.S. reviewers and non-U.S. reviewers differentially evaluate manuscripts that originate from within and outside of the United States. Link (1998) found that both categories of reviewers evaluate non-U.S. papers similarly. With respect to domestic papers, both categories of reviewers evaluated U.S.-authored papers more favorably (Link 1998). Still, we are unable to adjudicate whether U.S.-authored papers were more meritorious, resulting in their more favorable reception. That is, one cannot make an unequivocal conclusion about bias without measuring the quality of individual papers.

Opthof, Coronel, and Janse (2002) analyzed material submitted to *Cardiovascular Research* during a four year period (1997–2002). The researchers hypothesized that bias could relate to the country of origin of authors, reviewers, or both. The data for the analysis consisted of 3444 manuscripts linked to 10 countries around the globe. Analysis of reviewer scores showed that reviewers coming from the U.S. were significantly more likely to assign high priority to manuscripts, while reviewers from Japan, the United Kingdom, and Australia were significantly more likely to assign low priority scores. Authored manuscripts originating from the Netherlands, the United Kingdom, and the United States received significantly higher priority ratings, while authored manuscripts originating from Italy, Sweden, and "other countries" (a combination of countries where the number of authored manuscripts was too low to allow countries to be represented independently) scored significantly lower than average.

Because these ratings were merely suggestive of bias, the researchers differentiated reviewer-manuscript interactions in which the reviewer and author came from the same, and from different, countries. This analysis of "matched" and "non-matched" reviewer-manuscript interactions showed that manuscripts received significantly higher scores when reviewers and authors came from the same country. Opthof et al. (2002) also found a number of country-specific patterns. British and French reviewers, for example, assign significantly higher ratings to manuscripts from their own country compared with ratings from other countries. The overall patterns support a claim that geographical bias plays a role in the peer review process. This research suggests that part of the bias may stem from varying national and/or cultural standards that differentiate the practice of peer review. Still other research has produced results that fail to support or only moderately support claims of geographical bias (Primack & Marrs, 2008; Primack et al., 2009). The mixed results reflect the nascence of geography as an object of research on social particularism.

Any existence, however, of institutional, gender, and geographic bias prompts questions about whether and how reviewers are able to identify authors or their institutions in blind peer review. In blind review, author and institutional identity are

not supposed to be disclosed to reviewers. To better understand just how blind, blind peer review actually is, we turn lastly to research on author identification in the evaluation of scholarship.

Author Identification

While it is generally thought that double-blind review limits opportunities for social particularism because characteristics of authors are unknown to reviewers (Ware & Monkman, 2008), research lends only limited support of this view (see also Alam et al., 2011; Smith, Nixon, Bueschen, Venable, & Henry, 2002). Double-blind peer review may not eliminate social particularism because reviewers are able to identify authors or institutions even after names are redacted from manuscripts. A series of studies that examine reviewers' ability to identify authors and institutions help to understand the effectiveness of anonymization.

Several survey studies have indicated that reviewers often believe they can identify the authors of manuscripts. The studies show that between 50 and 80 % of reviewers (in psychology and physics) think they knew author identity (Adair, Carlon, & Sherman, 1981; Bradley, 1981; Ceci & Peters, 1984). There is little reason to suspect that reviewers in other fields would see themselves as less capable in this regard.

But research on the *accurate* identification of authors indicate that rates of success are lower but strikingly notable. The surveying of reviewers in the process of review has constituted the primary means by which researchers have investigated author and institutional identification. The earliest work on author identification, an analysis of 115 reviewers at the *Journal of Social Service Research*, found that one-third of reviewers were able to identify the authors of manuscripts they reviewed (Rosenblatt & Kirk, 1981). Building on this first round of research, Ceci and Peters (1984) conducted a similar study at six psychology journals that represented a broad range of specialties in the discipline. Of 146 reviewers participating in the study, 36 % accurately identified an author or co-author of the papers they reviewed.

Yankauer (1991) provided a questionnaire to reviewers at the *American Journal of Public Health* and found that 312 reviewers estimated that they could identify authors in 47 % of the cases; they were accurate 83 % of the time. In social services (Rosenblatt & Kirk, 1981), psychology (Ceci & Peters, 1984), medicine (Fisher, Freidman, & Strauss, 1994; McNutt, Evans, Fletcher, & Fletcher, 1990; Moossy & Moossy, 1985), and economics (Blank, 1991), studies have found reviewers are able to accurately identify authors more than one-third of the time (Weller, 2001). Further, Moossy and Moossy (1985) found that reviewers accurately identified 34 % of author institutions. This degree of failure of blind peer review to operate *as blind* is antithetical to the goal of disinterestedness. Author characteristics that are functionally irrelevant to merit are more readily considered in the assessment of manuscripts.

Simple editorial failures and author behavior have been flagged in research as among the key factors that make author and institutional identification possible. Self-citation in manuscripts has been credited as arguably the greatest giveaway. In Yankauer's (1991) study, author identification originated from self-citation 62 % of the time; personal knowledge of the research under review accounted for the balance. At the six psychology journals they studied, Ceci and Peters (1984) discovered that oversights in the editorial offices most often contributed to author identification. For example, some manuscripts were sent out containing title pages and acknowledgements to colleagues. When Ceci and Peters (1984) removed these cases from their analysis, author identification by reviewers dropped from 36 to 25 %, the lower figure still non-trivial.

Formal policies pertaining to the anonymization of authors are included in many, but not all, journals; it is widely treated as a professional norm that authors are to observe. Disclosure of identifying information is apparent even at journals with strict policies about anonymization. Editors at two radiology journals with double-blinded peer review policies analyzed 880 manuscripts without knowledge of authors' identity and institution and found that 34 % of the manuscripts (300) contained identifying information (Katz et al., 2002). Of the 300 manuscripts, the editors accurately identified the authors or institutions in 221 (74 %) of the cases, or 25 % of the overall manuscripts considered.

The source of identification is found in other behaviors. Authors include their initials in the body of the text; authors reference their own research; institutional identities are included in figures or in the text. Manuscripts have even been known to contain an author's name as a "key word" in the front matter of the submission. Thus, even when authors are presented with formal norms, if not also explicit directions, cues that signal author or institutional identity nevertheless end up in a substantial number of manuscripts operating under blind review.

Observations of author identification may be rendered in light of broader issues in the social organization of peer review. We earlier characterized the system of peer review as "strained" by virtue of the proliferation of journals and corresponding calls to participate in reviewing processes. The high and increasing volume of submissions likely creates conditions for error. Violations of policy would arguably decrease were editors and journal staff better positioned to vet the large number of submissions. A possible remedy is to increase the size of journal staff, including the number of editors at given journals. A version of this innovation has already occurred, wherein select journals now utilize a team of "area" or specialty deputy editors to aid in the processing of submissions. Still, this may serve as only a partial remedy to ensuring that blind review is, in fact, blind.

Some scholars may be punished for revealing their identity in manuscripts (e.g., a rejection on the basis of institutional bias against non-elite universities; a bias against younger scholars or graduate student authors; a desk-rejection). But bias can, of course, be expressed in negative and positive outcomes. That is, we have observed that the pressure to publish, not only at junior but also at senior levels, is now greater than ever. These conditions could create incentives to purposefully sig-

nal author identity; authors can believe that their reputation or their affiliation will work on behalf of a positive outcome. Thus, changing publication norms condition publication behavior, a part of which may inadvertently include an intensification of professional deviance. The practice of "blinding" is highly imperfect. The proliferation of publishing will likely make the occurrence of error more frequent. Thus the ground grows more fertile for bias. The explanation of these and related dysfunctional patterns discussed throughout the present section help to account for how the "sacred ideals" of peer review give way to "profane realities."

Broad Patterns and Directions for Future Research

In this chapter we assembled the central lines of research on peer review in science and scholarship. Our goal has been not only to take stock of what is known about peer review, but to also situate this body of knowledge sociologically and point to gaps that future research can address profitably. Three domains of work have been examined. First, we examined the objectives of peer review by discussing how it emerged, its relationship to the communication of knowledge and to the reward system of scholarship. Second, we examined how peer review is itself socially organized by accounting for its expansion and differentiation, its labor process, and its "invisible interaction." Third, we made apparent and examined dysfunctions that have arisen in contemporary peer review, including its reliability, the several forms of biases that work against its functional purposes, and a capacity to identify authors when most peer review processes are premised on being blind.

From the details of the research explored, broad patterns can be identified.

- The communication system of science is strained, and this strain is conducive to dysfunction, manifest, among many ways, in the time editors and reviewers dedicate to the evaluation of research, the quality of reviews, and the time it takes to complete reviews.
- There exists systemic disagreement among reviewers about the outcome of manuscripts under review. Generally, the peer review process of modern science and scholarship is unreliable, though this pattern may differ among fields of varying consensus.
- Although studies have yielded competing results on specific forms of bias (e.g., gender as opposed to institution), there is an undeniable presence of cognitive and social particularism in the operation of peer review. Richard Horton, editor of *The Lancet*, summarized the situation in stark terms:

> Editors and scientists alike insist on the pivotal importance of peer review. We portray peer review to the public as a quasi-sacred process that helps make science our most objective truth teller. But we know that the system of peer review is biased, unjust, unaccountable, incomplete, easily fixed, often insulting, usually ignorant, occasionally foolish, and frequently wrong (Horton, 2000, p. 149).

Extant research tends to provide support for many of these claims. But the body of research also makes clear that critics should be cautious about exaggerating the deleterious effects of peer review (Hargens and Herting 1991). We consider five gaps in research that identify ways to strengthen future research and extend knowledge about peer review in publication processes. In turn, the filling of these gaps may improve the practice of peer review.

First, there is need for methodological innovation. A major issue turns on the fact that only journal editors have complete information about the inner workings of peer review. This fact accounts for why a large portion of research is produced in the form of *post facto* reports by editors, of journal specific analyses, and of commentaries on peer review that rely more on illustrative cases than on data collected systematically for the purpose of addressing a research question. Some of the most groundbreaking research has been innovative because of the gains it realized through the use of deception. By fabricating papers, or by varying their characteristics across reviewers, Epstein (1990), Mahoney (1977), Peters and Ceci (1982), and Sokal (1996) enabled a more direct perspective into the operation of reliability and bias than inferred from, for example, inter-rater reliability scores or percentages of papers published by authors at elite versus non-elite universities. But with these innovations, and some of the unsavory findings on peer review that they exposed, came resistance, controversy, charges of misconduct, and movements to censure authors.

An important step forward methodologically consists of incorporating a fuller range of mechanisms in analyses. Taking the research on bias as an example, much of the work simply examines published research and reports differences in the representation of authors by gender, institutional affiliation, and nationality as supportive evidence. Several key mechanisms are missing to determine actual bias. Proxies that account for reviewer quality, time invested in the review, author characteristics, and manuscript characteristics would focus the lens on the existence of bias.

What is more, methodological progress ought to include more systematic analysis of rejected manuscripts. "Rejection without revision" constitutes the outcome of the majority of manuscript submissions, and thus research incorporating this category of outcomes is vital to understanding peer review processes comprehensively (Strang & Siler, 2015). Adopting mechanisms flagged above would cast light on the processes that send manuscripts in one direction as opposed to another—be it in first, second, or subsequent rounds of review, at one or more outlets.

A second gap exposes what comparatively little we know about peer review of publication in the humanities. Almost all of the peer review research draws upon data from the social, medical, and natural sciences. Substantial portions of academic faculty are based in the humanities (Schuster and Finkelstein 2006). Peer review is no less crucial to the building of these fields. Shatz (2004) has argued that the absence of research on peer review in the humanities is a result of disparate epistemological orientations to truth. According to this view, perceptions of truth vary more in the humanities than in the sciences, and thus it is more problematic to study peer review. Nevertheless, merit is demonstrated by evidence and argument, in comparative literature as in physics. Research on peer review in the humanities represents a virtually unexplored area of inquiry.

Third, there is a paucity of research on the peer review of books. Books are central to the production of knowledge in the humanities, the social sciences, and professional fields such as education and social work. To date, Powell's (1985) ethnography of scholarly publishing houses provides the most recent, systematic examination of decision-making in book publishing. A student of Coser, Powell's book is derivative of his advisor's master work on the same subject (Coser et al., 1982).

Other perspectives on the peer review of books can be gleaned from commentaries by university and commercial press editors (Appel 1994; Rowson 1994; Thatcher 1994), but these contributions say more about the coordinating roles of editors than about mechanisms of evaluation. The peer review of books is distinctive from journal publishing in several ways, chief among them the probability of manuscript evaluation. A majority of journal articles are initially rejected but ultimately accepted elsewhere. In book publishing, fields exhibit highly stringent standards for which houses count as high quality. This leaves authors of book manuscripts with comparatively fewer publication options. Because press editors receive high volumes of manuscripts, social networks between editors and authors take on a critical role (Powell, 1985). Relationships matter, and they matter more in publishing books than articles (where they are not supposed to matter at all). Editors seek to retain their "star" authors, meaning cumulative factors such as visibility and past accomplishments may operate alongside merit in the evaluation of manuscripts by well-established academics. The book world has changed dramatically since Powell's and Coser's key works. As at journals, competition for book publication has intensified. But where there is little question about the future of article publication, there are many questions about the future of academic monographs. Their future hinges on securing profitable manuscripts. How do editors get their books? How do they decide which to publish? On what conditions do they secure one type of reviewer as opposed to another? Why is there more editorial patience to receive the revisions on some work and not others? What are editors looking for, today, in reviews? These, and related questions, plow new ground in the world of peer review and book publishing.

A fourth gap that merits scholarly attention consists of cross-disciplinary analysis of peer review processes. We are in short supply of data on how peer review operates comparatively across fields. Instead, assumptions are made about peer review across sets of fields based on restrictive samples. Field consensus offers an analytic means to embark on such comparison (Braxton & Hargens, 1996). Is peer review systematically different in psychology than sociology, in educational psychology than leadership and policy studies? We have become comfortable in saying, often emphatically, that fields are different (Becher and Trowler 2001). Why should we be catholic about peer review? The opening section of this chapter underscored the normative premise that peer review is universalistic, not only in the criteria to be used to judge contributions, but also in its operation across what may be quite different worlds of science and scholarship. Comparative work would situate peer review processes in a matrix of consensus, codification, and the production of knowledge across fields.

Finally, it is worthwhile to investigate the ways by which outlet hierarchies might condition reviews. One might speculate that lower quality outlets generate comparatively lower quality reviews, and vice versa at higher quality outlets. Powell (1985) has shown that standards of quality are applied differentially. "Authors were called upon to uphold the standards that they themselves had helped establish" (Powell, 1985, p. 156). This suggests that standards are learned, and presumes that only after publishing in high status outlets are scholars able to legitimately conduct reviews for such outlets.

The more general implication is that the standards applied by reviewers vary by outlet status. This constitutes its own form of particularism; where there is particularism, there are significant implications for the production and functioning of science and scholarship. A study in the field of ecology demonstrated that as scientists publish in high impact journals, they more often recommend rejection (Aarssen et al., 2009). In filling this gap, researchers can show how processes of peer review—varied as they may be across outlets of publication—create subcultures of quality. These micro-worlds impinge directly on, indeed create, the stocks of knowledge that define fields. If the stocks vary in value, we may be called once more to question the application of merit on the one hand, and the incidence of bias on the other.

Since its inception in the mid-1600s, but especially across its subsequent maturation in the twentieth and twenty-first centuries, we have learned that the questioning of process differentiates the sacred ideals from the profane realities of peer review. Peer review operates as both the "linchpin of science" and a "lightning rod" for scientists, given both its centrality to, and deep imperfections in, constituting academic work. To these ends, peer review implicates—by a propagating sequence of linkages—the authenticity of larger structures in which academe is organized. If so much depends on peer review, and if its operation is so riddled particularistically, then as much can purportedly be said of academic reward systems and the academic profession itself. Systems of reward and professional organization are, of their own accord, fundamental to academe. They, too, are predicated in cognitive rationalism. We expect rewards to be fair, and professions, upstanding. But therein lie propagated disjunctures between ideals and realities. Macro systems of reward and professions cannot perform universalistically and meritocratically when reviews of scholarly and scientific performance are so compromised in their micro dynamics. The study of peer review thus makes clear that there is much at stake. The present review has identified areas of peer review processes in critical need of attention. The substance and order of all of academic life will be its beneficiary.

References

Aarssen, L. W., Lortie, C. J., Budden, A. E., Koricheva, J., Leimu, R., & Tregenza, T. (2009). Does publication in top-tier journals affect reviewer behavior? *Plos One, 4*(7), e6283. doi:10.1371/journal.pone.0006283.

Abbott, A. (2011). Personal communication to second author, annual board meeting of the *American Journal of Sociology*, Atlanta, Georgia.

Abramowitz, S. I., Gomes, B., & Abramowitz, C. V. (1975). Publish or politic: Referee bias in manuscript review. *Journal of Applied Social Psychology, 5*(3), 187–2000.

Adair, R. K., Carlon, H. R., & Sherman, C. (1981). Anonymous refereeing. *Physics Today, 34*(6), 13–15.

Alam, M., Kim, N.A., Havey, J., Rademaker, A., Ratner, D., Tregre, B.,…Coleman, W.P. (2011). Blinded vs. unblended peer review of manuscripts submitted to a dermatology journal: A randomized multi-rater study. *British Journal of Dermatology*, *165*(3), 563–567.

Alperts, B. (2013). Impact factor distortions. *Science, 340*(6134), 787.

Amrein, K., Langmann, A., Fahrleitner-Pammer, A., Pieber, T. R., & Zollner-Schwetz, I. (2011). Women underrepresented on editorial boards of 60 major medical journals. *Gender Medicine, 8*(6), 378–387.

Aper, J. P., & Fry, J. E. (2003). Post-tenure review at graduate institutions in the United States. *Journal of Higher Education, 74*(3), 241–260.

Appel, C. S. (1994). University press editing and publishing. In R. J. Simon & J. J. Fyfe (Eds.), *Editors as gatekeepers: Getting published in the social sciences* (pp. 179–194). Lanham, MD: Rowman & Littlefield Publishers.

Bacon, F. (1620). *Novum Organum: Or, true suggestions for the interpretation of nature*. London, UK: Routledge and Sons.

Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental article in science*. Madison, WI: University of Wisconsin Press.

Becher, T., & Trowler, P. R. (2001). *Academic tribes and territories*. Buckingham, UK: Open University Press.

Becker, H. S. (1984). *Art worlds*. Berkeley, CA: University of California Press.

Ben-David, J. (1965). The scientific role: The conditions of its establishment in Europe. *Minerva, 4*(1), 15–54.

Bernard, H. R. (1980). Report from the editor. *Human Organization, 39*(4), 366–369.

Bess, J. L. (1988). *Collegiality and bureaucracy in the modern university: The influence of information and power on decision-making structures*. New York: Teachers College Press.

Beveridge, W. I. B. (1950). *The art of scientific investigation*. New York: Vintage Books.

Biagioli, M. (2002). From book censorship to academic peer review. *Emergences: Journal for the Study of Media & Composite Cultures, 12*(1), 11–45.

Björk, B. C., Roos, A., & Lauri, M. (2009). Scientific journal publishing: Yearly volume and open access availability. *Information Research*, *14*(1). http://InformationR.net/ir/14-1/paper391.html

Björk, B., Welling, P., Laakso, M., Majlender, P., Hedlund, T., & Guõnason, G. (2010). Open access to the scientific literature: Situation 2009. *PLoS ONE, 5*(6), e11273. doi:10.137/journal.pone.0011273.

Björk, B. C., & Solomon, D. (2013). The publishing delay in scholarly peer-reviewed journals. *Journal of Infometrics, 7*(4), 914–923.

Blank, R. M. (1991). The effects of double-blind versus single-blind reviewing: Experimental evidence from the *American Economic Review*. *American Economic Review, 81*(5), 1041–1067.

Bohannon, J. (2013). Who's afraid of peer review? *Science, 342*(6154), 60–65.

Bornmann, L., & Daniel, H. (2005). Selection of research fellowship recipients by committee peer review: Reliability, fairness and predictive validity of board of trustees' decisions. *Scientometrics, 63*(2), 297–320.

Bornmann, L., Mutz, R., & Daniel, H. (2010a). A reliability-generalization study of journal peer reviews: A multi-level meta-analysis of inter-rater-reliability and its determinants. *PloS One, 5*(12), e14331. doi:10.1371/journalpone.0014331.

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. Cambridge, MA: Harvard University Press.

Bradley, J. V. (1981). Pernicious publication practices. *Bulletin of the Psychonomic Society, 18*(1), 31–34.

Braxton, J. M., & Hargens, L. L. (1996). Variation among academic disciplines: Analytical frameworks and research. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. XI, pp. 1–46). New York, NY: Agathon.

Budden, A. E., Tregenza, T., Aarssen, L. W., Koricheva, J., Leimu, R., & Lortie, C. J. (2008). Double-blind review favours increased representation of female authors. *TRENDS in Ecology and Evolution, 23*(1), 4–6.

Burnham, J. C. (1990). The evolution of editorial peer review. *Journal of the American Medical Association, 263*(10), 1323–1329.

Calhoun, C. (2010). On Merton's legacy and contemporary sociology. In R. K. Merton (Ed.), *Sociology of science and sociology as science* (pp. 1–32). New York, NY: Columbia University Press.

Campanario, J. M. (1993). Consolation for the scientist: Sometimes it is hard to publish papers that are later highly-cited. *Social Studies of Science, 23*, 342–362.

Campanario, J. M. (1998). Peer review for journals as it stands today—Part 1. *Science Communication, 19*(1), 181–211.

Ceci, S. J., & Peters, D. (1984). How blind is blind peer review? *American Psychologist, 39*(12), 1491–1492.

Ceci, S. J., & Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences, 108*(8), 3157–3162.

Chase, J. M. (1970). Normative criteria for scientific publication. *American Sociologist, 5*(3), 262–265.

Cho, A. H., Johnson, S. A., Schuman, C. E., Adler, J. M., Gonzalez, O., Graves, S. J., et al. (2014). Women are underrepresented on the editorial boards of journals in environmental biology and natural resource management. *PeerJ, 2*, e542. doi:10.7717/peerj.542.

Chubin, D. E., & Hackett, E. J. (1990). *Peerless science: Peer review and U.S. science policy*. Stony Brook, NY: State University Press of New York.

Cicchetti, D. V. (1991). The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation. *Behavioral and Brain Sciences, 14*(1), 119–186.

Cicchetti, D. V., & Eron, L. D. (1979). The reliability of manuscript reviewing for the *Journal of Abnormal Psychology*. *Proceedings of the American Statistical Association, 22*, 596–600.

Cleary, F. R., & Edwards, D. J. (1960). The origins of the contributors to the A.E.R. during the 'fifties. *The American Economic Review, 50*(5), 1011–1014.

Cole, S. (1992). *Making science: Between nature and reality*. Cambridge, MA: Harvard University Press.

Cole, S., & Cole, J. R. (1973). *Social Stratification in Science*. Chicago, IL: University of Chicago Press.

Cole, S., Cole, J. R., & Simon, G. (1981). Chance and consensus in peer review. *Science, 214*(4523), 881–886.

Cole, S., Rubin, L., & Cole, J. R. (1978). *Peer review in the National Science Foundation: Phase one of a study*. Washington, DC: National Academy of Sciences.

Coser, L. A., Kadushin, C., & Powell, W. W. (1982). *Books: The culture and commerce of publishing*. Chicago, IL: University of Chicago Press.

Coursol, A., & Wagner, E. E. (1986). Effect of positive findings on submission and acceptance rates: A note on meta-analysis bias. *Professional Psychology: Research and Practice, 17*(2), 136–137.

Crane, D. (1965). Scientists at major and minor universities: A study of productivity and recognition. *American Sociological Review, 30*, 699–714.

Crane, D. (1967). The gatekeepers of science: Some factors affecting the selection of articles for scientific journals. *The American Sociologist, 2*(4), 195–201.

Cummings, L., Frost, P. J., & Vakil, T. F. (1985). The manuscript review process: A view from the inside on coaches, critics, and special cases. In L. L. Cummings & P. J. Frost (Eds.), *Publishing in the organizational sciences* (pp. 469–508). Homewood, IL: Irwin.

Dar, R. (1987). Another look at Meehl, Lakatos, and the scientific practices of psychologists. *American Psychologist, 42*(2), 145–151.

De Vries, R., Anderson, M. S., & Martinson, B. C. (2006). Normal misbehavior: Scientists talk about the ethics of research. *Journal of Empirical Research on Human Research Ethics, 1*(1), 43–50.

Dewitt, N., & Turner, R. (2001). Bad peer reviewers. *Nature, 413*(6852), 93.

Dressler, A. J. (1970). Nobel prizes: 1970 awards—Physics. *Science, 170*, 604–606.

Edwards, L.N., & Ferber, M.A. (1986). Journal reviewing practices and the progress of women in the economics profession: Is there a relationship? *Newsletter of the Committee on the Status of Women in the Economics Profession*, 1–7.

Emerson, G. B., Winston, J. W., Wolf, F., Heckman, J., Brand, R., & Leopold, S. S. (2010). Testing for the presence of positive-outcome bias in peer review: A randomized controlled trial. *Archives of Internal Medicine, 170*(21), 1934–1939.

Epstein, W. (1990). Confirmational response bias among social work journals. *Science, Technology, and Human Values, 15*(1), 9–38.

Ernst, E., & Kienbacher, T. (1991). Chauvinism. *Nature, 353*(6336), 560.

Ernst, E., Resch, K. L., & Uher, E. M. (1992). Reviewer bias. *Annals of Internal Medicine, 116*(1), 958.

Fairweather, J. S. (2002). The ultimate faculty evaluation: Promotion and tenure decisions. *New Directions for Institutional Research, 114*, 97–108.

Fanelli, D. (2010). "Positive" results increase down the hierarchy of the sciences. *PLos One, 5*(4), e10068. doi:10.1371/journal.pone.0010068.

Ferber, M. A., & Teiman, M. (1980). Are women economists at a disadvantage in publishing journal articles? *Eastern Economic Journal, 6*(3–4), 189–193.

Finke, R. A. (1990). Recommendations for contemporary editorial practices. *American Psychologist, 45*(5), 669–670.

Fisher, M., Freidman, S. B., & Strauss, B. (1994). The effects of blinding on acceptance of research papers by peer review. *JAMA, 272*(2), 143–148.

Fox, M. F. (1994). Scientific misconduct and editorial peer review processes. *Journal of Higher Education, 65*(3), 298–309.

Friedkin, N. (1983). Horizons of observability and limits of informal control in organizations. *Social Forces, 62*(1), 54–77.

Fyfe, J. J. (1994). Cops and robbers in academe: Editing *Justice Quarterly*. In R. Simons & J. J. Fyfe (Eds.), *Editors as gatekeepers: Getting published in the social sciences* (pp. 59–72). Boston, MA: Rowman and Littlefield.

Garfield, E. (1955). Citation indexes to science: A new dimension in documentation through association of ideas. *Science, 122*(3159), 108–111.

Garfunkel, J. M., Ulshen, M. H., Hamrick, H. J., & Lawson, E. (1994). Effect of institutional prestige on reviewers' recommendations and editorial decisions. *JAMA, 272*(2), 137–138.

General Accounting Office. (1994). *Peer review: Reforms needed to ensure fairness in federal agency grant selection*. Washington, DC: General Accounting Office.

Geiger, R. L. (1986). *To advance knowledge: The growth of American research universities, 1900–1940*. New York, NY: Oxford University Press.

Gilbert, J. R., Williams, E. S., & Lundberg, G. D. (1994). Is there a gender bias in JAMA's review process? *JAMA, 272*(2), 137–138.

Gillespie Jr., G. W., Chubin, D. E., & Kurzon, G. M. (1985). Experience with NIH peer review: Researchers' cynicism and desire for change. *Science, Technology, and Human Values, 10*(3), 44–54.

Goldberg, P. (1968). Are women prejudiced against women? *Trans-Action, 5*(5), 28–30.

Goodrich, D. W. (1945). An analysis of manuscripts received by the editors of the American Sociological Review from May 1, 1944 to September 1, 1945. *American Sociological Review, 10*(6), 716–725.

Greaves, S., Scott, J., Clarke, M., Miller, L., Hannay, T., Thomas, A., & Campbell, P. (2006). Overview: *Nature*'s trial of open peer review. *Nature*. doi:10.1038/nature05535.

Greenwald, A. G. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin, 82*, 1–20.

Grouse, L. D. (1981). The Ingelfinger rule. *Journal of the American Medical Association, 245*(4), 375–376.

Guest, A. M. (1994). Gatekeeping among the demographers. In R. Simons & J. J. Fyfe (Eds.), *Editors as gatekeepers: Getting published in the social sciences* (pp. 85–106). Boston, MA: Rowman and Littlefield.

Guetzkow, J., Lamont, M., & Mallard, G. (2004). What is originality in the humanities and social sciences? *American Sociological Review, 69*(2), 190–212.

Gustin, B. H. (1973). Charisma, recognition, and the motivation of scientists. *American Journal of Sociology, 78*(5), 1119–1134.

Hall, R. A., & Hall, M. B. (Eds.) (1966). *The correspondence of Henry Oldenburg*, Vol. 2. Madison, WI: University of Wisconsin Press.

Hamermesh, D. S. (1994). Facts and myths about refereeing. *The Journal of Economic Perspectives, 8*(1), 153–163.

Hargens, L. L. (1988). Scholarly consensus and journal rejection rates. *American Sociological Review, 53*, 139–151.

Hargens, L. L., & Herting, J. R. (1990). Neglected considerations in the analysis of agreement among journal referees. *Scientometrics, 19*, 91–106.

Hersen, M., & Miller, D. J. (1992). Future directions: A modest proposal. In D. J. Miller & M. Hersen (Eds.), *Research fraud in the behavioral and biomedical sciences* (pp. 225–244). New York, NY: John Wiley.

Harnad, S. (1982). *Peer commentary on peer review: A case study in scientific quality control*. New York, NY: Cambridge University Press.

Hearn, J. C., & Anderson, M. S. (2002). Conflict in academic departments: An analysis of disputes over faculty promotion and tenure. *Research in Higher Education, 43*(5), 503–529.

Hermanowicz, J. C. (1998). *The stars are not enough: Scientists—Their passions and professions*. Chicago, IL: University of Chicago Press.

Hermanowicz, J. C. (2009). *Lives in science: How institutions shape academic careers*. Chicago, IL: University of Chicago Press.

Hermanowicz, J. C. (2016a). The proliferation of publishing: Economic rationality and ritualized productivity in a neoliberal era. *American Sociologist, 47*, 174–191.

Hermanowicz, J. C. (2016b). Universities, academic careers, and the valorization of 'shiny things.'. *Research in the Sociology of Organizations, 46*, 303–328.

Herrera, A. J. (1999). Language bias discredits the peer-review system. *Nature, 297*(6719), 467.

Hess, D. J. (1997). *Science studies: An advanced introduction*. New York, NY: New York University Press.

Hirschauer, S. (2010). Editorial judgments: A praxeology of 'voting' in peer review. *Social Studies of Science, 40*(1), 71–103.

Hood, J. (1985). The lone scholar myth. In M. F. Fox (Ed.), *Scholarly writing and publishing* (pp. 111–124). Boulder, CO: Westview Press.

Horton, R. (2000). Genetically modified food: consternation, confusion, and crack-up. *Medical Journal of Australia, 177*, 148–149.

Jacobs, J. A. (2013). *In defense of disciplines: Interdisciplinarity and specialization in the research university*. Chicago,IL: University of Chicago Press.

Jacobs, J. A., & Winslow, S. E. (2004). Overworked faculty: Job stresses and family demands. *The Annals of the American Academy of Political and Social Science, 596*(1), 104–129.

Jauch, L. R., & Wall, J. L. (1989). What they do when they get your manuscript: A survey of *Academy of Management* reviewer practices. *Academy of Management Journal, 32*(1), 157–173.

Judson, H. F. (1994). Structural transformations of the sciences and the end of peer review. *Journal of the American Medical Association, 272*(2), 92–94.

Juhasz, S., Calvert, E., Jackson, T., Kronick, D., & Shipman, J. (1975). Acceptance and rejection of manuscripts. *IEEE Transactions on Professional Communication, 18*(3), 177–185.

Katz, D. S., Proto, A. V., & Olmsted, W. W. (2002). Incidence and nature of unblinding by authors: Our experience at two radiology journals with double-blinded peer review policies. *American Journal of Roentgenology, 179*(6), 1415–1417.

Kerr, S., Tolliver, J., & Petree, D. (1977). Manuscript characteristics which influence acceptance for management and social science journals. *Academy of Management Journal, 20*(1), 132–141.

Knight, J. (2003). Negative results: Null and void. *Nature, 422*(6932), 554–555.

Knorr-Cetina, K. (1981). *The manufacture of knowledge: An essay on the constructivist and contextual nature of science*. Oxford,UK: Pergamon Press.

Kravitz, D. J., & Baker, C. I. (2011). Toward a new model of scientific publishing: Discussion and a proposal. *Frontiers in Computational Neuroscience, 5*(55), 1–12.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

Lamont, M. (2009). *How professors think: Inside the curious world of academic judgment*. Cambridge, MA: Harvard University Press.

Langfeldt, L. (2001). The decision-making constraints and processes of grant peer review, and their effects on the review outcome. *Social Studies of Science, 31*(6), 820–841.

Lawrence, J. H., Celis, S., & Ott, M. (2014). Is the tenure process fair? What faculty think. *Journal of Higher Education, 85*(2), 155–192.

Leahey, E., & Cain, C. L. (2013). Straight from the source: Accounting for scientific success. *Social Studies of Science, 43*(6), 927–951.

Lee, C. J., & Schunn, C. D. (2011). Social biases and solutions for procedural objectivity. *Hypatia: A Journal of Feminist Philosophy, 26*(2), 352–373.

Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology, 64*(1), 2–17.

Levenson, H., Burford, B., Bonno, B., & Davis, L. (1975). Are women still prejudiced against women? A replication and extension of Goldberg's study. *Journal of Psychology, 89*, 67–71.

Liebert, R. J. (1976). Productivity, favor, and grants among scholars. *American Journal of Sociology, 82*(3), 664–673.

Link, A. M. (1998). US and non-US submissions. *Journal of the American Medical Association, 280*(3), 246–247.

Lloyd, M. E. (1990). Gender factors in reviewer recommendations for manuscript publication. *Journal of Applied Behavior Analysis, 23*(4), 539–543.

Long, J. S., Allison, P. D., & McGinnis, R. (1993). Rank advancement in academic careers: Sex differences and the effects of productivity. *American Sociological Review, 58*, 703–722.

Long, J. S., & Fox, M. F. (1995). Scientific careers: Universalism and particularism. *Annual Review of Sociology, 21*, 45–71.

Loonen, M. P. J., Hage, J. J., & Kon, M. (2005). Who benefits from peer review? An analysis of the outcome of 100 requests for review by *Plastic and Reconstructive Surgery*. *Plastic and Reconstructive Surgery, 116*(5), 1461–1472.

Mahoney, M. J. (1977). Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research, 1*(2), 161–175.

McAfee, R. P. (2014). Edifying editors. In M. Szenberg & L. Ramrattan (Eds.), *Secrets of Economics Editors* (pp. 33–44). Cambridge, MA: MIT Press.

McLachlan, J. C. (2010). Integrative medicine and the point of credulity. *BMJ, 341*, c6979.

McNutt, R. A., Evans, A. T., Fletcher, R. H., & Fletcher, S. W. (1990). The effects of blinding on the quality of review. *JAMA, 263*(10), 1371–1376.

Merton, R. K. (1973a). The normative structure of science. In N. W. Storer (Ed.), *The sociology of science: Theoretical and empirical investigations* (pp. 267–278). Chicago, IL: University of Chicago Press (Article originally published in 1942).

Merton, R. K. (1973b). Priorities in scientific discovery. In N. W. Storer (Ed.), *The sociology of science: Theoretical and empirical investigations* (pp. 286–324). Chicago, IL: University of Chicago Press (Article originally published in 1957).

Merton, R. K. (1973c). The Matthew effect in science. In N. W. Storer (Ed.), *The sociology of science: Theoretical and empirical investigations* (pp. 439–459). Chicago, IL: University of Chicago Press (Article originally published in 1968).

Merton, R. K., & Zuckerman, H. (1973). Age, aging, and the age structure in science. In N. W. Storer (Ed.), *The sociology of science: Theoretical and empirical investigations* (pp. 497–559). Chicago, IL: University of Chicago Press (Article originally published in 1972).

Moossy, J., & Moossy, Y. R. (1985). Anonymous authors, anonymous referees: An editorial explanation. *Journal of Neuropathology and Experimental Neurology, 44*(3), 225–228.

Morton, H. C., & Price, A. J. (1986). The ACLS survey of scholars: Views on publications, computers, libraries. *Scholarly Communication, 5*, 1–16.

Mulkay, M. (1980). Interpretation and the use of rules: The case of the norms of science. In T. F. Gieryn (Ed.), *Science and social structure: A Festschrift for Robert K. Merton* (pp. 111–125). New York, NY: New York Academy of Sciences.

National Science Board. (2010). *Science and engineering indicators 2010*. Arlington, VA: National Science Foundation.

Olson, C. M., Rennie, D., Cook, D., Dickersin, K., Flanagin, A., Hogan, J., … Pace, B. (2002). Publication bias in editorial decision making. *Journal of the American Medical Association, 287*(21), 2825–2828.

O'Meara, K. A. (2004). Beliefs about post-tenure review: The influence of autonomy, collegiality, career stage, and institutional context. *Journal of Higher Education, 75*(2), 178–202.

Opthof, T., Coronel, R., & Janse, M. (2002). The significance of the peer review process against the background of bias: Priority ratings of reviewers and editors and the prediction of citation, the role of geographical bias. *Cardiovascular Research, 56*(3), 339–346.

Orr, R. & Kassab, J. (1965). *Peer group judgments on scientific merit: Editorial refereeing*. Presentation to the Congress of the International Federation for Documentation, Washington, DC. October 15.

Paludi, M. A., & Bauer, W. D. (1983). Goldberg revisited: What's in an author's name? *Sex Roles, 9*(3), 387–390.

Patriquin, L., Bensimon, E. M., Polkinghorne, D. E., Bauman, G., Bleza, M. G., Oliverez, P. M., & Soto, M. (2003). Posttenure review: The disparity between intent and implementation. *Review of Higher Education, 26*(3), 275–297.

Perna, L. W. (2001). Sex and race differences in faculty tenure and promotion. *Research in Higher Education, 42*(5), 541–567.

Perna, L. W. (2003). Studying faculty salary equity: A review of theoretical and methodological approaches. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 18, pp. 323–388).

Perna, L. W. (2005). Sex differences in faculty tenure and promotion: The contribution of family ties. *Research in Higher Education, 46*(3), 277–307.

Peters, D. P., & Ceci, S. J. (1982). Peer-review practices of psychological journals: The fate of accepted, published articles, submitted again. *Behavioral and Brain Sciences, 5*(2), 187–195.

Polanyi, M. (1946). *Science, faith, and society*. Chicago, IL: University of Chicago Press.

Powell, W. W. (1985). *Getting into print: The decision-making process in scholarly publishing*. Chicago, IL: University of Chicago Press.

Power, M. (1997). *The audit society: Rituals of verification*. Oxford, UK: Oxford University Press.

Price, D. J. (1975). *Science since Babylon*. New Haven, CT: Yale University Press.

Primack, R. B., Ellwood, E., Miller-Rushing, A. J., Marrs, R., & Mulligan, A. (2009). Do gender, nationality, or academic age affect review decisions? An analysis of submission to the journal *Biological Conservation*. *Biological Conservation, 142*, 2415–2418.

Primack, R. B., & Marrs, R. (2008). Bias in the review process. *Biological Conservation, 141*(12), 2919–2920.

Ramos-Alvarez, M. M., Moreno-Fernandez, M. M., Valdes-Conroy, B., & Catena, A. (2008). Criteria of the peer review process for publication of experimental and quasi-experimental research in psychology: A guide for creating research papers. *International Journal of Clinical and Health Psychology, 8*(3), 751–764.

Ripp, A. (1985). Peer review is alive and well in the United States. *Science, Technology, and Human Values, 10*(3), 82–86.

Rosenblatt, A., & Kirk, S. A. (1981). Recognition of authors in blind review of manuscripts. *Journal of Social Science Research, 3*(4), 383–394.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638–641.

Rowson, R. C. (1994). A formula for successful scholarly publishing: Policy-oriented research and the humanities. In R. J. Simon & J. J. Fyfe (Eds.), *Editors as gatekeepers: Getting published in the social sciences* (pp. 195–208). Lanham, MD: Rowman & Littlefield Publishers.

Roy, R. (1985). Funding science: The real defects of peer review and an alternative to it. *Science, Technology, and Human Values, 10*(3), 73–81.

Schuster, J. H., & Finkelstein, M. J. (2006). *The American faculty: The restructuring of academic work and careers*. Baltimore, MD: Johns Hopkins University Press.

Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *BMJ, 314*(7079), 498–502.

Shadish Jr., W. R. (1989). The perception and evaluation of quality in science. In B. Gholson, W. R. Shadish Jr., R. A. Neimeyer, & A. C. Houts (Eds.), *Psychology of science: Contributions to metascience* (pp. 383–426). Cambridge, UK: Cambridge University Press.

Shadish Jr., W. R., Doherty, M., & Montgomery, L. M. (1989). How many studies in the file drawer? An estimate from the family/marital psychotherapy literature. *Clinical Psychology Review, 9*(5), 589–603.

Shatz, D. (2004). *Peer review: A critical inquiry*. Lanham, MD: Rowman & Littlefield.

Shils, E. (1997). *The order of learning: Essays on the contemporary university*. New Brunswick, NJ: Transaction.

Siler, K., Lee, K., & Bero, L. (2015). Measuring the effectiveness of scientific gatekeeping. *Proceedings of the National Academy of Sciences, 112*(2), 360–365.

Silverman, R. (1988). Peer judgment: An ideal typification. *Knowledge: Creation, Diffusion, Utilization, 9*, 362–382.

Simon, R. J., Bakanic, V., & McPhail, C. (1986). Who complains to journal editors and what happens. *Sociological Inquiry, 56*(2), 259–271.

Smart, R. G. (1964). The importance of negative results in psychological research. *Canadian Psychology, 5*, 225–232.

Smigel, E. O., & Ross, H. L. (1970). Factors in the editorial decision. *The American Sociologist, 5*(1), 19–21.

Smith, J. A., Nixon, R., Bueschen, A. J., Venable, D. D., & Henry, H. H. (2002). The impact of blinded versus unblended abstract review on scientific program content. *Journal of Urology, 168*(5), 2123–2125.

Sokal, A. D. (1996). A physicists experiments with cultural studies. *Lingua Franca, 6*(4), 62–64.

Starbuck, W. H. (2003). Turning lessons into lemonade: Where is the value in peer reviews? *Journal of Management Inquiry, 12*(4), 344–351.

Sterling, T. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—Or vice versa. *Journal of the American Statistical Association, 54*(285), 30–34.

Stossel, T. P. (1985). Refinement in biomedical communication: A case study. *Science, Technology, & Human Values, 10*(3), 39–43.

Strang, D., & Siler, K. (2015). Revising as reframing: Original submissions versus published papers in *Administrative Science Quarterly*, 2005–2009. *Sociological Theory, 33*(1), 71–96.

Sugimoto, C. R., Larivière, V., Ni, C., & Cronin, B. (2013). Journal acceptance rates: a cross-disciplinary analysis of variability and relationships with journal measures. *Journal of Informatics, 7*(4), 897–906.

Tenopir, C., & King, D. W. (2009). The growth of journals publishing. In B. Cope & A. Phillips (Eds.), *The future of the academic journal* (pp. 105–124). Oxford, UK: Chandos Publishing.

Teplitskiy, M. (2015). Frame search and re-search: How quantitative sociological articles change during peer review. *American Sociologist*. doi:10.1008/s12108-015-9288-3.

Thatcher, S. G. (1994). Listbuilding at university presses. In R. J. Simon & J. J. Fyfe (Eds.), *Editors as gatekeepers: Getting published in the social sciences* (pp. 209–258). Lanham, MD: Rowman & Littlefield Publishers.

Travis, G., & Collins, H. M. (1991). New light on old boys: cognitive and institutional particularism in the peer review system. *Science, Technology, & Human Values, 16*(3), 322–341.

Tuchman, G. (2009). *Wannabe U: Inside the Corporate University*. Chicago, IL: University of Chicago Press.

Valkonen, L., & Brooks, J. (2011). Gender balance in *Cortex* acceptance rates. *Cortex, 47*, 763–770.

Wanderer, J. J. (1966). Academic origins of contributors to the "American Sociological Review", 1955–1965. *The American Sociologist, 1*(5), 241–243.

Ward, C. (1981). Prejudice against women: Who, when, why? *Sex Roles, 7*(2), 163–171.

Ware, M. & Monkman, M. (2008). *Peer review in scholarly journals: Perspective of the scholarly community—An international study*. Publishing Research Consortium.

Waters, M. (1989). Collegiality, bureaucratization, and professionalization: A Weberian Analysis. *American Journal of Sociology, 94*(5), 945–972.

Webb, T. J., O'Hara, B., & Freckleton, R. P. (2008). Does double-blind review benefit female authors? *Trends in Ecology and Evolution, 23*(7), 351–353.

Weisse, A. B. (1986). Say it isn't no: Positive thinking and the publication of medical research. *Hospital Practice, 21*(3), 23–25.

Weller, A. C. (2001). *Editorial peer review: Its strengths and weaknesses*. Medford, NJ: American Society for Information Science and Technology.

Wessely, S. (1996). What do we know about peer review? *Psychological Medicine, 26*, 883–886.

Wessely, S., Brugha, T., Cowen, P., Smith, L., & Paykel, E. (1996). Do authors know who refereed their paper? A questionnaire survey. *BMJ, 313*(7066), 1185.

Whittaker, R. J. (2008). Journal review and gender equality: A critical comment on Budden et al. *Trends in Ecology and Evolution, 23*(9), 478–479.

Wilhite, A. W., & Fong, E. A. (2012). Coercive citation in academic publishing. *Science, 335*(6068), 542–543.

Wood, M., & Johnsrud, L. (2005). Post-tenure review: What matters to faculty. *Review of Higher Education, 28*(3), 393–420.

Yankauer, A. (1990). Who are the peer reviewers and how much do they review? *JAMA, 263*(10), 1338–1340.

Yankauer, A. (1991). How blind is blind review? *American Journal of Public Health, 81*(7), 843–845.

Youn, T. I. K., & Price, T. M. (2009). Learning from the experience of others: The evolution of faculty tenure and promotion rules in comprehensive institutions. *Journal of Higher Education, 80*(2), 204–237.

Ziman, J. M. (1968). *Public knowledge: An essay concerning the social dimension of science*. Cambridge, UK: Cambridge University Press.

Zuckerman, H. (1977). *Scientific elite: Nobel laureates in the United States*. New York, NY: The Free Press.

Zuckerman, H., & Merton, R. K. (1971). Patterns of evaluation in science: Institutionalization, structure and the functions of the referee system. *Minerva, 9*, 66–100.

Zuckerman, H. (1988). Sociology of science. In N. J. Smelser (Ed.), *Handbook of sociology* (pp. 511–574). Newbury Park, CA: Sage.

David R. Johnson is an Assistant Professor of Higher Education at the University of Nevada-Reno. His research examines how universities are reconfiguring their social contract with society and the implications of changing social organization of science for stratification and inequality in the academic profession. He is the author of *A Fractured Profession: Commercialism and Conflict in Academic Science* (Johns Hopkins University Press, 2017). Other examples of his research on higher education include “Ethical Ambiguity in Science” (*Science and Engineering Ethics*), “Narratives of Science Outreach in Elite Contexts of Academic Science” (*Science Communication*), and “Technological Change and Professional Control in the Professoriate” (*Science, Technology, and Human Values*).

Joseph C. Hermanowicz is professor of sociology and adjunct professor in the Institute of Higher Education at the University of Georgia. His interests center on the academic profession, academic careers, the operation of reward systems in organizations, and qualitative research methodology. He is the author, most recently, of “Honor in the Academic Profession: How Professors Want to Be Remembered by Colleagues” (*Journal of Higher Education*), “How Professors View Their Graduate Education” (*Higher Education*), and “The Culture of Mediocrity” (*Minerva*). In addition, he is the author of *Lives in Science: How Institutions Affect Academic Careers* (University of Chicago Press, 2009), *The Stars Are Not Enough: Scientists—Their Passions and Professions* (University of Chicago Press, 1998), and *College Attrition at American Research Universities: Comparative Case Studies* (Agathon, 2003) and editor of *The A merican Academic Profession: Transformation in Contemporary Higher Education* (Johns Hopkins University Press, 2011).