Peer Review: The View from Social Studies of Science

Christina K Pikas
January 29, 2008
for STS Independent Study

1. Introduction

Peer review is central to science, and it is under attack. Publishers argue that open access journals will end peer review and socialize science.[1] Cases of scientific fraud have exposed media and public misperceptions of the purposes of peer review as well as weaknesses in the current system.[2] For these reasons, peer review is an ideal case study for the application and discussion of social studies of science. This paper reviews studies of peer review and provides approaches to understanding peer review in its social context through various lenses of social studies of science. The paper concludes with some suggestions to address some of peer review's perceived failings and to take advantage of some of the benefits of modern information and communication technologies (ICTs).

According to Merton (1973) science is "a set of characteristic methods by means of which knowledge is certified, a stock of accumulated knowledge stemming from the application of these methods, a set of cultural values and mores governing the activities termed scientific, or any combination of the foregoing" (p. 268). Reports of completed work in journals are an important method used to certify knowledge and to build and preserve the stock of scientific information. Peer reviewed journal articles are the end product of basic scientific work (Latour & Woolgar, 1986).

Scientists must publish to establish priority, to extend knowledge, to assert their position in the scholarly community, and to advance their position in their institution (Polanyi, 2000). After the experiments, analysis, and write up are complete, the paper has to pass the final major hurdle prior to publication: peer review. Peer review, frequently called refereeing or editorial peer review, is the use of knowledgeable readers on the topic to evaluate the merits of the work and suggest publication, corrections and re-submission, or rejection (Weller, 2001). Peer review is also used to award grants and to provide scientific oversight to the regulatory process, but these types of peer review will not be discussed here.

Many articles and books on the subject of peer review start with recounting the history of peer review back to the early days of the Royal Society (see, for example Zuckerman & Merton, 1971). For the purposes of describing how peer review currently works and how it might be updated to address concerns and take advantage of new technologies, I consider only more recent applications of peer review within the scholarly publishing system.

This article starts with by describing the purpose of peer review and continues by detailing how peer review works, the place of peer review in society, ways of judging peer review, and problems with peer review. Then there is a brief synopsis of recent research. This article concludes by suggesting some changes to the peer review system grounded in the understanding of the social place of peer review developed through the article.

---

[1] *Socialize* in this case is used as a contrast to a "free market" system. See the Association of American Publishers' PRISM Coalition press releases available from: http://www.prismcoalition.org/index.htm (accessed 1/24/2008).

[2] See discussions of the Hwang case in two important journals: Peer review and fraud.(2006). *Nature*, *444*(7122), 971-972. DOI:10.1038/444971b and Couzin, J. (2006). Breakthrough of The Year: Breakdown of the Year: Scientific fraud. *Science, 314*(5807), 1853. DOI:10.1126/science.314.5807.1853

2. Purpose of Peer Review

Peer review has many purposes, some explicit and some implicit purposes developed over time. Primarily, peer review is used to decide what to publish in journals with limited space. It is also used to improve the quality of the research and the statistics, as well as to certify scientific knowledge, build community, and to train scientists.

2.1. Deciding What to Publish

Polanyi (2000) provides a practical guide to how scientists determine scientific merit. The primary criteria are plausibility, scientific value, and originality. Plausibility speaks to whether the conclusions are sound based on the current state of knowledge and whether the topic meshes with how the reviewer views the phenomenon. The Duhem-Quine Thesis states that a theory can never be conclusively tested in isolation because it depends on an entire framework or belief system (Sismondo, 2004). The scientific data do not completely determine the results; interpretation is required (Golinski, 1998).

How the new article fits into the reviewer's mental models and "webs of belief" will greatly impact his or her decision on its plausibility (Collins, 1985). Latour and Woolgar (1986) describe various ways to provide evidence of this plausibility including through referencing previous work. Further, less well supported or new findings are couched in modalities to lessen their weight. Tables, photographs, and other included evidence also lend credibility.

Polanyi (2000) states that scientific value is a function of three coefficients:

- accuracy
- systematic importance
- the intrinsic interest of the subject matter (p. 5).

The overarching guidelines for determining scientific merit are translated into pragmatic guidelines for reviewers. These guidelines frequently ask the reviewers to answer variations of the following questions:

- does the article fit, is it appropriate for the venue?
- is it good science?
- is it original, non-trivial, and interesting?
- is it well done?

2.1.1 How well the paper matches the scope

One of the first decisions in peer review is whether the topic, the treatment, the research methods, and aspects of the write-up such as length and style match the venue. The editor or editorial board might make decisions on scope without sending the article out to reviewers (See more on process below; Zuckerman & Merton, 1971). Reviewers may be asked to suggest a more appropriate venue (see instructions from *Angewandte Chemie* available at: http://www.wiley.com/WileyCDA/Section/id-301837.html)

2.1.2 How to determine good science

Social studies of science tells us a great deal about how to philosophically determine good science (or even if the theory is scientific at all) and about how practicing scientists actually

judge science. Scientists judge the merit of scientific theories several ways. Popper (1968) states that a scientific theory is better if it

- has greater content
- is logically stronger
- has greater explanatory and predictive power
- can be more severely tested (p.217).

2.1.3 How to determine if the article is original, non-trivial, and interesting

There is a tension between plausibility and originality. If the article is to be considered plausible, it must build on and refer to the certified knowledge of the field both through the research methods selected and through referencing; however, a truly original article might propose new techniques, use different terminology, and rely less on other articles (Kuhn, 1996; Latour & Woolgar, 1986; Polanyi, 2000). If the article is using new techniques that are unfamiliar to the reviewers or are from a revolutionary paradigm, it will be difficult for the reviewer to accept the work (Kuhn, 1996). Nevertheless, the article must make a contribution to the state of knowledge, and must be important to the system and interesting.

An aspect of originality is that the article must be non-trivial. That is, the article must make new knowledge claims that are interesting and increase the knowledge base of the field. In some cases, articles replicating earlier studies, or providing evidence for theories or explanations proposed in earlier work are accepted[3].

2.2. Improve Quality

Reviewers are frequently encouraged to make suggestions to help the author improve the paper. For example, *Nature*'s peer review guidelines state

> The primary purpose of the review is to provide the editors with the information needed to reach a decision. The review should also instruct the authors on how they can strengthen their paper to the point where it may be acceptable. As far as possible, a negative review should explain to the authors the weaknesses of their manuscript, so that rejected authors can understand the basis for the decision and see in broad terms what needs to be done to improve the manuscript for publication elsewhere. This is secondary to the other functions, however, and referees should not feel obliged to provide detailed, constructive advice to authors of papers that do not meet the criteria for the journal (as outlined in the letter from the editor when asking for the review).
> (http://www.nature.com/authors/editorial_policies/peer_review.html)

A recent survey of authors of peer reviewed scientific journal authors found that 64% reported that peer review of their last article identified scientific errors. Many more authors reported that peer review improved their paper's readability, presentation, and inclusion of references (Mark Ware Consulting, 2008).

Peer review also acts as a filter, to lessen the burden of scientists who are trying to keep up in their fields.

---

[3] This depends on the goals of the journal. Compare, for example, the guidelines for *Physical Review Letters* with those from *Angewandte Chemie*.

2.3. Build Community

Reviewing scientific articles is time consuming and is not compensated. The reviewers participate to get access to scientific articles prior to publication, to contribute to their scientific community, and for prestige (Mark Ware Consulting, 2008). Zuckerman and Merton (1971) suggest that referees in fields with poorly developed informal communication networks or that are rapidly developing will have an advantage over scientists who do not review in learning about new work.

3. How Peer Review Works

The general process currently used by most scientific journals follows the pattern shown in figure 1.

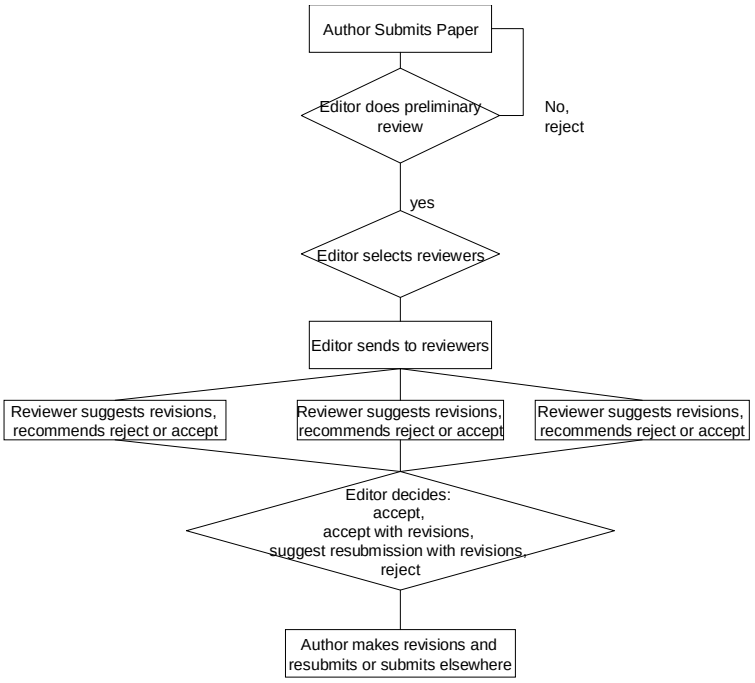

Figure 1: General Peer Review Process

When the author submits an article to a journal using an online system, an editor checks the article to make sure it is at least reporting science, is somewhat in scope, and is complete. Zuckerman and Merton (1971) found that for journals with high rejection rates editors may reject half of the total number of rejections without submitting the papers to reviewers and a journal with very low rejection rates accepted about 80% of the total acceptances without using reviewers. The editor, or in some cases the editorial board or publisher representative, selects two to three reviewers and sends them the review and a deadline for responses (Mark Ware Consulting, 2008). The package might contain a checklist, guidelines, a form, and confidentiality rules. In biomedicine, a special statistics reviewer might be selected in addition to the content reviewers.

Reviewers read the article, and return a form with their recommendation to accept, accept with revisions, require revision and resubmission for further review, or rejection. Generally, the form will have room for additional suggestions for improvements. The editor or editorial team

reviews and then makes a decision based on the strength of the comments from the various reviewers. The editor then notifies the author who can make revisions and resubmit, withdraw the article and possibly resubmit elsewhere, or who can appeal the decision.

Each step of this process has complicating factors. Selection of reviewers, the reviewers' processes, and actions taken by the editorial board are selected for more detailed discussion.

3.1. The Difficulty of Selecting Reviewers

Editorial teams and publishers must decide to whom to send articles for review. The reviewers must have sufficient knowledge of the topic to adequately judge the article; have time to complete the review; have a good track record in performing reviews in a timely, through, and fair manner; and perhaps should not be a direct competitor for priority or funding.

Not selecting direct competitors as reviewers may seem counter-intuitive based on the Mertonian norms of science (universalism, communism, disinterestedness, and organized skepticism) (1973). In contrast to these institutional norms, Polanyi and others describe the affective, intellectual, and personal commitment to theories and priority claims that could prove too tempting in reviewing competitors' work (cited in Mitroff, 1974).

Collins (1985) defines a core set as the group of allies and critics who are interested parties in a scientific controversy. It is the members of this set who submit articles on a topic and who are perhaps best suited to review new work. The existence of a single ally in a reviewer or editorial position may lead to the success of a theory.

On the other hand, Polanyi (2000) describes overlapping neighborhoods in science such that while scientists can admittedly exercise competent judgment only over a small part of science, they can usually judge an area adjoining their own special studies that is broad enough to include some fields on which other scientists have specialized.

3.2. Anonymity

Reviews may be double blinded, single blinded, signed, or open. In double blinded reviews, the article's authors' names are removed and the names of the reviewers are not provided to the author. In single blinded reviews, the authors' names are provided to the reviewers, but the reviewers' names are not provided to the author. In both of these cases, the reviewer might choose to sign his or her review. More recently, open review has come to mean that the article will be made available for anyone to review, in a web forum.

4. Peer Review in Society

Peer review of journals has social implications both within science and in interfaces between science and non-scientists.

4.1. How Scientists Use Peer Review

Scientists have to rely on and trust other scientists' work in place of direct experiential knowledge (Shapin, 1995). Peer review and post-publication proxies for quality such as citedness can help scientists trust papers.

4.2. How Non-Scientists Use Peer Review

Peer reviewed articles are very important in establishing the "epistemic authority" of science (Sismondo, 2004). Epistemic authority is the right to make knowledge claims that will be accepted. Science has authority because it is self-policing. In Jasanoff's (1987) words "the process of peer review… reinforces the position of science as an autonomous social institution requiring no external control."

Regulators, courts, journalists, and members of the public rely on the journals' peer review systems as an imprimatur (Jasanoff, 1987, Sismondo, 2004). Scientific experts are qualified through their peer reviewed articles as well as their institutional affiliation.

5. Problems With Peer Review

In the Mertonian idealized view, disinterestedness is an institutional norm of science (Merton, 1973). Fraud in science is minimized through the "exacting scrutiny of fellow experts" and the "ultimate accountability of scientists to their compeers" (p.276); yet there have been a number of well-publicized cases of fraud that were not caught by the pre-publication editorial or peer review process and other cases of plagiarism or other misbehavior that happened in the peer review process.

5.1. Conservative Bias

Peer review is inherently conservative. Reviewers judge plausibility based on what they know and expect (Polanyi, 2000). revolutionary ideas are difficult, if not impossible, to communicate to scientists who are in the normal paradigm due to incommensurability (Kuhn, 1996).

Further, scientists are affectively, personally, and intellectually committed to their own theoretical viewpoint. Mitroff(1974) found many differences between the impersonal character of science described in Merton's earlier articles (see Merton, 1973) and the personal character of science revealed through his study of elite scientists of the Apollo program. The scientists were not disinterested, but were passionately committed to their theoretical view. Even a reviewer who is consciously trying to be fair and open-minded will be less likely to have as positive views of articles based on a competing theoretical framework.

5.2. Gender, Underrepresented Groups, and Stratification

There are suggestions that women, members of underrepresented groups, and scientists from less prestigious institutions are less likely to win grants, are less likely to have papers accepted for publication, and get less support for their research. J.R. Cole and S. Cole (1973) reviewed data on the research output of women scientists matched with male counterparts and found that women were less productive, even when regressed for prestige of institution, marital status, and number of children. Part of this difference could be attributable to bias in peer review but more information is needed.

Merton (1973) noted that greater recognition goes to scientists who are established and who have won awards. The "Matthew Effect," as he called it, lead him to propose that scientific contributions from higher ranking scientists will have greater visibility and can therefore make a greater contribution to scientific knowledge. Some of this is explained by taking into account the fact that higher ranking scientists will be able to get better institutional support, better research

assistants, and will be able to be more selective in problems to study as well as articles to publish.

Two recent studies showed that double blinded review increased the acceptance rate of articles written by women and authors from less prestigious institutions (Budden et al., 2008; Ross et al., 2006); however, there have been many other studies that indicated that it is very difficult if not impossible to keep the author's identity secret (see discussion in Jefferson, Rudin, Brodney Folse, & Davidoff, 2007). Members of a small research community will know what work topics other scientists are working on and will recognize the author's self-citations in the reference list. Van Rooyen's randomized controlled trial of blinded vs. unblinded found no difference in the quality of the review, acceptance rate, or time taken to review (cited in Jefferson et al, 2007, p1998).

Zuckerman and Merton's (1971) study of the effects of age and rank on article acceptance found that younger authors' papers were more likely to be accepted and higher ranking authors' papers were more likely to be accepted. Their data did not support hypotheses regarding the relative ranks of the authors and reviewers.

5.3. Bad Behavior

Despite the idealized view of science as impersonal, universal, and disinterested, there are very real inducements to violate the norms. Some examples of bad behavior by reviewers mentioned in the literature are:

- reviewers steal or plagiarize content from articles they are reviewing
- reviewers try to stifle competing paradigms
- reviewers delay their review until their own competing work is published to get priority for their claims (Zuckerman & Merton,1971)

Guidelines to reviewers often indicate that the reviewers should treat articles they are reviewing as confidential and that they are to declare conflicts of interest.

5.4. Not Enough Information

In most cases neither the reviewer nor the reader of the final work has access to the data or enough information on the methods to replicate the experiment or detect misconduct (Overbeke, 1999). Collins (1985) has named this the experimenter's regress: it is difficult to judge the scientific results without the judging the competency of the experiment and it is difficult to judge the experiment without reviewing the results. The way out of the regress is to judge the work using other features beyond the results and the method. Besides not having the actual data, the reviewer can not be expected to determine if the work is a duplication of previous work or is plagiarism unless his or her work is directly involved[4].

For some very specialized articles on the frontiers of science, there might not be a fully qualified reviewer who can adequately judge the work (Zuckerman & Merton,1971).

[4] There are some horror stories of reviewers being sent work that clearly plagiarizes their own work.

5.5. Delays and Expense

The peer review process is very costly in the editor's time and the reviewers' time. In a recent survey, editors reported that the average time from submission to final acceptance is 130 days (Mark Ware Consulting, 2008). Some journals have letters or rapid communications sections that are published after less strenuous review and somewhat more quickly (see, for example, *Angewandte Chemie*, and *Physical Review*).

6. Recent Research on Peer Review

Many research, theoretical, and opinion articles have been published and continue to be published on peer review, particularly in the biomedical sciences. Figure 2 provides some idea of the number of publications[5]. There are also several books reviewing the topic and as well as books making political statements about some aspect of the process. One of the more comprehensive reviews is Weller's published in 2001. Weller reviewed 1,500 articles and books.

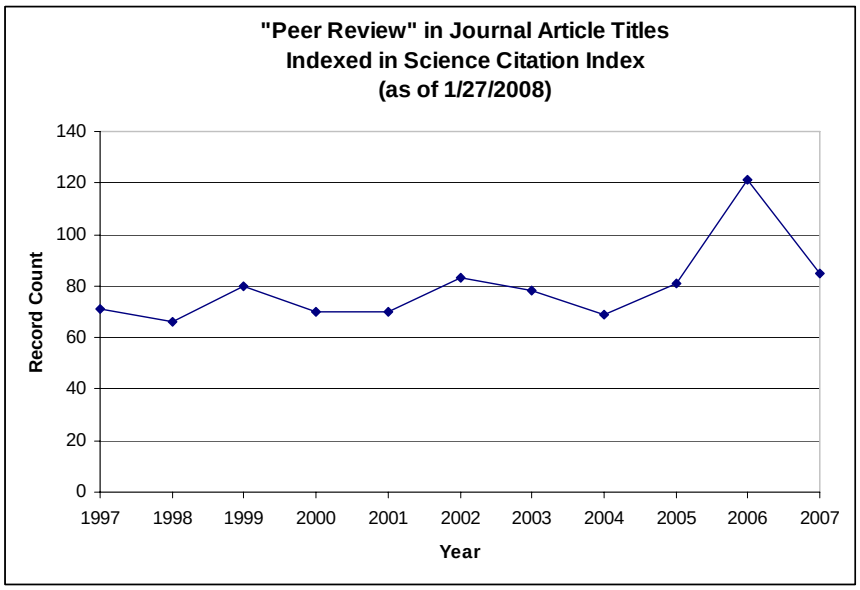

Figure 2

More recently, Jefferson, Rudin, Brodney Folse, and Davidoff (2007) completed a systematic review of editorial peer review in biomedical research. They included only 28 "prospective or retrospective comparative studies with two or more comparison groups generated by random or other appropriate methods" (p.1). They concluded that there is not enough empirical evidence that peer review ensures quality in research.

In research areas outside of biomedicine, Scott (2007) questions whether peer reviewers should also consider issues of social relevance. Other recent studies tested the impact of blinding the reviews, training the reviewers, and using electronic formats.

[5] Note: some of these articles are not about the scientific journal peer review process, but by inspection most are. More thorough searching in abstracting and indexing services still must use natural language terms because few thesauri in the sciences seem to include terms for peer review. A notable exception is MeSH and this points even more to the importance of discussions of peer review in the biomedical discourse.

6.1. How to Judge Peer Review

Judging the quality, utility, or efficiency of peer review in a generic sense or for a particular venue has been the focus of several recent reviews, most of which end with calls for more rigorous research. Many ways of measuring have been used.

- Quality of Outcome. Several studies used panels to rate the quality of articles before and after peer review, or to measure the quality of the comments provided by the reviewers. One study inserted several errors into an article to see how many were caught by the various reviewers (cited in Jefferson et al, 2007).
- Acceptance/rejection rate
- If the articles eventually were published elsewhere
- Fairness. Several efforts have been made to judge the fairness of peer review by randomized trial. Other studies have done statistical analyses retrospectively.
- Inter-reviewer agreement. Several studies have shown low levels of agreement between reviewers (Overbeke, 1999).
- Authors' and editors' satisfaction

7. Suggested Changes to the Peer Review System

The peer review system has been in the news from time to time and efforts have been made to update it. Some things that have been tried recently include double blinding, signed peer review, open peer review, and post publication peer review. Suggestions of professionalizing peer reviewers in a peer review bureau have also been made. As seen above, double blinding has had limited success, depending on the number of scientists active in the research area. Efforts by *Nature* and *Medical Journal of Australia* have had limited success in open peer review due to low participation rates (Overview: Natures Trial of Peer Review, 2008; Jefferson et al, 2007). Post-publication review can be used as an addition to the traditional peer review, or in lieu of traditional peer review.

A very new (January 1, 2008) development is the Neuroscience Peer Review Consortium (http://nprc.incf.org/). A group of neuroscience journals have agreed to forward reviews from one journal to the next upon author request. The member journals hope that this will minimize some redundancy and mitigate some of the burden on the pool of reviewers.

7.1. Review the Reviewer

Several discussions of peer review indicated authors decline to review papers for reasons other than conflicts of interest. In fact, the Ware study found that authors declined between 15-25% of reviewer requests and the most frequent reason was that they were too busy (Mark Ware Consulting, 2008). The same study found that the participants reviewed 8 papers per year on average and on average spent 8.5 hours on each review. Further, some reviewer reports are much better than others. With the expense and the time spent in carefully preparing reports, it is unfortunate that these reports are only seen by a very few people and that credit for being a reviewer is binary, not weighted for the value of the contribution.

I propose first that reviewers be allowed to designate some of their comments public, and these comments be published alongside the articles in the online version of the journal. As it stands, reviewers can designate part of their comments for the editors only and not to be seen by the authors. This would add another area on the report. These comments could be signed or not,

depending on the agreement of the editors and the reviewer. This way, careful, thoughtful comments could gain the reviewer social capital.

I propose also, that reviewer databases be shared across similar journals so that certain reviewers are not overly burdened, and potential reviewers who submit many articles but do not carry their share of the load can be encouraged to review new papers. Adding in information on the quality of the reviews might be helpful, but might be troublesome if an author felt he or she was being unfairly discriminated against. A simple way to add information on the quality of the reviews is to do a social software rating system. Journal editors might feel that their reviewer databases provide a competitive advantage for their journal and so might be unwilling to share.

7.2. Modernize the Process

Thoughtful summary comments and rankings are important, but additional feedback is lost in translating any notes or comments or highlighting the reviewer might have made on the copy he or she read. I propose that reviewers be given an online version that allows them to markup, circle, annotate, and interact more fully with the text. Adobe Acrobat markup tools might be adequate. The Institute for the Future of the Book's Comment Press provides an open source software option (http://www.futureofthebook.org/commentpress/). Summary comments and ratings would still be necessary.

8. Conclusions

Despite newsworthy attacks on peer review, it is a system that scientists trust and that works for their purposes. Close study of the peer review process exposes the social system of science, and this does include bad or at least self-interested behavior in some cases, but also tremendous amounts of high-value work that is uncompensated. Peer review is not broken, but could be made more efficient to continue to keep pace with the expansion of science. The system is not equipped to identify fraud, duplicate publication, and plagiarism; but the reports in the media seem to indicate that the public thinks it is. More information for the public in science popularizations might help.

References

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. DOI:10.1016/j.tree.2007.07.008

Cole, J. R., & Cole, S. (1973). *Social stratification in science*. Chicago: University of Chicago Press.

Collins, H. M. (1985). The scientist in the network. *Changing order: Replication and induction in scientific practice* (pp. 129-157). Beverly Hills, CA: Sage Publications.

Golinski, J. (1998). *Making natural knowledge: Constructivism and the history of science*. Cambridge, U.K.: Cambridge University Press.

Jasanoff, S. S. (1987). Contested boundaries in policy-relevant science. *Social Studies of Science, 17*(2), 195-230.

Jefferson, T., Rudin, M., Brodney Folse, S., & Davidoff, F. (2007). Editorial peer review for improving the quality of reports of biomedical studies. *Cochrane database of systematic reviews* (). Chichester, UK: John Wiley & Sons.

Kuhn, T. S. (1996). *The structure of scientific revolutions* (3rd ed.). Chicago, IL: University of Chicago Press.

Latour, B., & Woolgar, S. (1986). *Laboratory life :The construction of scientific facts*. Princeton, N.J.: Princeton University Press.

Mark Ware Consulting. (2008). *Peer review in scholarly journals: Perspective of the scholarly community – an international study*. Bristol, UK: Author.

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

Mitroff, I. I. (1974). Norms and counter-norms in a select group of the apollo moon scientists: A case study of the ambivalence of scientists. *American Sociological Review, 39*(4), 579-595.

Overview: Nature's trial of open peer review. Retrieved 1/29/2008, 2008, from http://www.nature.com/nature/peerreview/debate/nature05535.html

Overbeke, J. (1999). The state of the evidence: What we know and what we don't know about journal peer review. In F. Godlee, & T. Jefferson (Eds.), *Peer review in health sciences* (pp. 32-44). London: BMJ Books.

Polanyi, M. (2000). The republic of science: Its political and economic theory. *Minerva: A Review of Science, Learning & Policy, 38*(1), 1-21.

Popper, K. R. (1968). Truth, rationality, and the growth of scientific knowledge. *Conjectures and refutations: The growth of scientific knowledge* (pp. 215-250). New York: Harper & Row.

Ross, J. S., Gross, C. P., Desai, M. M., Hong, Y., Grant, A. O., Daniels, S. R., et al. (2006). Effect of blinded peer review on abstract acceptance. *JAMA: Journal of the American Medical Association, 295*(14), 1675-1680. DOI:10.1001/jama.295.14.1675

Scott, A. (2007). Peer review and the relevance of science. *Futures, 39*(7), 827-845. DOI:10.1016/j.futures.2006.12.009

Shapin, S. (1995). Here and everywhere: Sociology of scientific knowledge. *Annual Review of Sociology, 21*(1), 289-321.

Sismondo, S. (2004). *An introduction to science and technology studies*. Malden, MA: Blackwell.

Weller, A. C. (2001). *Editorial peer review: Its strengths and weaknesses*. Medford, N.J.: Information Today.

Zuckerman, H., & Merton, R. K. (1971). Patterns of evaluation in science: Institutionalization, structure and functions of the referee system. *Minerva 9*(1), 66-100.