[unreviewed preprint submitted to New Review of Information Networking]

Peer Review: recent experience and future directions

Introduction

It is not difficult to find statements in the literature decrying the “burden” of peer review or the crisis in the peer review system. For example:

> “*The peer review system is breaking down and will soon be in crisis: increasing numbers of submitted manuscripts mean that demand for reviews is outstripping supply*” (Fox & Petchey 2010)
>
> "*The peer-review system ... the foundation on which scientific advance is based, is near to breaking point*” (Franklin 2010)

Is this really the case? This article reviews the evidence and takes a look at recent developments in the field.

It certainly seems likely that interest in peer review has never been higher. The academic study of peer review as a topic in its own right is a recent one, pioneered by medical editors at BMJ and JAMA and dating approximately from the first International Peer Review Congress in 1989. There has been a steady growth in academic papers on peer review, with about 200 papers a year currently appearing in PubMed. The result of this interest is that there is now a sizeable evidence base covering many areas of peer review, although ironically this evidence base appears little known and often ignored by journal editors.

Peer review today

The scale of the problem

Let us start with some basic facts about the peer-reviewed journal system. There are estimated to be about [26,000] peer-reviewed, scholarly journals as of 2010 (of which about 9500 are included in the Thomson Reuters ISI database). The number of journals has grown by about 3% per year. (Ware & Mabe 2009).

These journals in total publish about 1.5 million articles per year (of which roughly a million appear in journals in the ISI database), a number that has also consistently grown at about 2.5–3% per year. According to Björk (2009) the average ISI-listed journal publishes 112 articles per year and the average non-ISI-listed journal 26 articles, giving a weighted average of 57 articles per year.

The number of articles published is highly correlated to (and presumably driven by) the number of active researchers, which in turn depends on levels of R&D funding. Notwithstanding shorter term economic cycles, therefore, over the long term it is very likely that the number of articles produced will continue to grow in line with the longterm trend. The link to economic output is also clear in the growth of output from East Asia, particularly China and India, whose shares of the global article output have grown dramatically in recent years. In some fields, particularly in the hard sciences and engineering, many journals now receive more submissions from China than any other country. This increasing globalisation of science has implications for peer review.

The typical peer-reviewed journal will send submitted articles to review by 2–3 reviewers per paper (the overall average is 2.3). The most commonly used system is single-blind review, in which the author's name is known to the reviewer but not vice versa. Double-blind review, in which reviewers' and authors' identities are both concealed, is less common, particularly in science but is more common in humanities and social science. Open peer review can mean the opposite of double-blind, that authors' and reviewers' identities are both known to each other (and sometimes publicly disclosed) but discussion is complicated by the fact that it is also used to describe other approaches, such as where the reviewers remain anonymous but their reports are published.

The large majority of peer reviewing is now organised using online tracking systems; a 2007 survey (Ware 2008) found that over three-quarters of editors reported using these systems, with higher proportions in the life sciences (85%) but somewhat lower take-up in the humanities and social sciences (51%).

Rejection rates at peer-reviewed journals vary widely, from over 90% to under 10%, with the average rejection rate being estimated to be about 50%. The path of rejection and resubmission taken by many papers means they undergo more than one cycle of review.

The overall average (median) time spent by reviewers per articles is about 5 hours (mean 8.5 hours), though this figure will vary by discipline and also by experience (more active reviewers report spending less time per review than less frequent reviewers).

The costs of peer review are made up of the time spent by researchers in performing the review; this is a non-cash cost, as they are not paid directly to do this. There are also the costs of publishers in administering the process, and in payments to Editors and (infrequently) to reviewers. The non-cash element was estimated by the economic consultancy CEPA (RIN/CEPA 2008) to represent the large majority of total peer review costs, at about £2 billion or roughly £1200 per paper.

Perceptions and misperceptions

We have a good snapshot of the perceptions (and arguably in some cases, misperceptions) of the research community with respect to peer review from two large-scale surveys conducted by Ware (2008) and Sense About Science (2009). These surveys were international and covered all disciplines. Researchers surveyed were those who had recently published in journals covered by the ISI databases, which may give some bias towards the experience of higher quality or more international or English-language journals.

To start with, it was clear from these surveys and others that researchers strongly supported peer review, seeing it as necessary and as important to effective control of communication in science. As a respondent to a RIN study on Web 2.0 use by researchers said: "Anything that isn’t peer reviewed ... is worthless" (RIN 2010). The same study reported that peer-reviewed journal articles were considered “very important” information sources by 92% of researchers compared to 4% giving that rating to “un-refereed articles”.

Support for peer review is not the same as uncritical support for all its aspects or being satisfied with its operation in practice. We found that some two-thirds of researchers described themselves as satisfied and only 12% dissatisfied with the operation of peer review.

It is also the case that researchers may say one thing and do another: for example, saying that peer review is critical while at the same time making increasing use of other measures (such as author affiliation, funding sources, etc.) to evaluate and use non-peer-reviewed materials such as preprints (Armbruster 2007).

The strong level of support for peer review is matched by widespread belief in its effectiveness (although as we will see below, not all of this belief in effectiveness is warranted by the available evidence). About 90% of researchers overall thought the main area of effectiveness for peer review was in improving the published paper.

On this point, researchers might point to their own experience. Asked to consider specifically their last published paper, researchers overwhelmingly (90%) said that it had been improved by peer review. Areas of improvement included the substantive, such scientific and statistical content and missing or inaccurate references, as well as presentational and language-related issues.

Substantial percentages of researchers (around 80% in the Sense About Science survey) thought that peer review was effective in dealing with plagiarism, fraud and other types of academic misconduct. This is more surprising, given some high-profile cases and the lack of evidence in this area.

One interesting finding was that researchers expressed a clear preference for double-blind over other types of peer review (single-blind, open or post-publication) and saw double-blind as being more effective. This was despite single-blind review being much more commonly experienced than double-blind; the preference for double-blind was stronger among those who had experienced it. In fact, a majority of respondents in the humanities and social sciences (where double-blind review is more common) thought that single-blind review was not effective. The reasons given for preferring double-blind review were principally to do with objectiveness and fairness, reducing the scope for various kinds of bias.

Reviewers' motivations

Why is it that researchers are prepared to put in the time for peer review, given that it is rarely counted towards professional advancement?

The most frequent reason given in the surveys was that it was simply part of the obligations of their professional life; for example over 90% said they reviewed to play their part as a member of the academic/research community. As most reviewers are also authors, there is a reciprocal element: researchers say they are willing to review because they appreciate the benefits they receive as authors. Researchers also cite motivations such as being able to see work ahead of publication and enjoying being able to help improve the paper.

Researchers in the surveys were much less likely to cite self-interested than these more altruistic reasons, but some did agree that things like enhancing their professional reputations, increasing the chance of being appointed to an editorial board or currying favour with an Editor played a part.

Overall, payments to reviewers are rare (though token payments do exist at some journals, for instance in economics and at the BMJ). In surveys, researchers split fairly evenly on the desirability of paying reviewers in principle but support for payment drops away when respondents are asked to consider the costs of payment being met through subscriptions or (especially) by author payments.

Some of the support for payment may derive from increased debate over the role of commercial publishers or it may simply reflect a desire for the contribution to be acknowledged. There was substantial support, for instance for reviewers to be acknowledged by name (for instance in an annual list of reviewers).

Recognition for reviewers appears something of a two-edged sword. On the one hand, researchers appear to value some acknowledgement or recognition of their contribution (Ware 2008; Groves 2006; Aspinall 2009). On the other hand, the majority of reviewers seem keen to not to have their name published alongside the article, or otherwise be publicly identified as the reviewer of a particular article. We shall cover this in more detail in the discussion of open peer review.

The main areas of discontent identified in these researcher surveys are (perhaps predictably): concern as authors at the length of time taken by the process; some concern at the burdens imposed by reviewing commitments (although in absolute terms the reported workloads did not appear very onerous, with active reviewers reporting an average of 14 reviews per year at 5 hours each); and concerns about bias and lack of fairness.

Other author complaints (e.g. Nickerson 2005) about peer review include poor quality reviewing, e.g. superficial, vague, unsupported; inappropriately self-serving feedback, e.g. attempting to increase citations to the reviewer's own work; or less frequently outright misconduct (deliberately delaying work, plagiarising ideas, etc.) One also hears complaints of "criticism creep" whereby the reviewer requests so many detailed changes they start to become in effect a co-author.

Benefits

Researcher perceptions are valuable but do not give the whole picture. We can identify a number of possible ways in which peer review might benefit the academic and wider community.

Peer review could provide a quality assurance process for improving the quality of research studies (as distinct from improving the submitted manuscript prior to publication). Although some see this as one of its purposes, this sets a very high bar for peer review and at present there is little evidence to show its effectiveness in this way (e.g. Jefferson et al 2007).

On the other hand, there seems reasonable albeit self-reported evidence (e.g. the surveys reported above) that peer review does improve published papers. It is likely to do this in two ways: first, the existence of a hurdle (and the associated competition for space in journals) encourages authors to be self-critical prior to submission; and second, through the feedback and revision process inherent in peer review itself.

Peer review also acts as a filter, to the benefit of readers. For professional researchers, the most important aspect of this filtering is not just the fact that peer review has taken place, but the basis it provides for the stratification of journals by perceived quality: peer review is the process that routes better articles to better and/or most appropriate journals. While there is some debate regarding whether journal brands are still the most efficient or effective way to filter the literature (see the discussion of post-publication review below), there seems little doubt that at present they remain very important to researchers as readers as well as as authors (e.g. Tenopir et al 2010).

Peer review can also act as a kind of "seal of approval" for readers, distinguishing the peer-reviewed from the non-peer-reviewed literature. This is probably more important for lay readers and journalists, particularly in relation to pseudo-science or particularly poor work. There are, however, problems with promoting peer review as a kind of badge of quality assurance: as mentioned above, there is little evidence of its effectiveness in this regard, and it is not really designed to detect errors, misconduct or outright fraud. It is also the case that almost any genuine academic manuscript, however weak, can find a journal to publish it if the author is persistent enough.

It's worth noting that in this picture authors as well as readers benefit from peer review, from feedback that helps them improve their papers. Proposals to replace (rather than supplement) pre-publication review with post-publication systems (e.g. comments, ratings and other metrics) would lose this benefit.

Critiques of peer review

Peer review is of course not without its critics. Most criticism falls into one or more of these categories: validity, effectiveness, efficiency or fairness. With the growth of research into peer review itself, much of this criticism is now grounded in good evidence.

Perhaps the most important criticism of peer review is that it is ineffective, that it doesn't work. Richard Smith, the previous Editor of the BMJ (and current board member of PLoS) has been one of the most vocal and consistent critics, describing it as a flawed system and a waste of reviewers' time (e.g. Smith 2006; Smith 2009).

The Cochrane Collaboration has reviewed the evidence for the effectiveness of peer review as a quality assurance tool and concluded that "little empirical evidence is available to support the use of editorial peer review as a mechanism to ensure quality of biomedical research" (Jefferson 2007), although they go on to concede that there are methodological problems in studying peer review and point out that the absence of evidence is not the same as evidence of absence. They also do find some evidence that editorial peer review does make papers more readable and improve the quality of reporting.

What does it mean to say that peer review is unreliable? If it were an effective test of the quality (or just publishability) of a manuscript, you would expect reviewers' opinion to be consistent and agree with each other. In practice, a host of studies have shown this is not the case. One study showed that the probability of reviewers agreeing to be only slightly greater than chance; to obtain a statistically reliable result, editors would need to obtain reviews from six or more reviewers, rather than just two (Rothwell and Martyn 2000). Bornmann et al (2010) conducted a meta-analysis of the available studies and concluded that the inter-rater reliability was limited. Bornmann & Daniel (2010) also showed that inter-rater reliability was not improved in public peer review (as practised by *Atmospheric Chemistry & Physics*) compared to traditional closed peer review. In an example of the consequences of this unreliability, Bornmann & Daniel (2009) demonstrated that in a substantial fraction of cases (about a quarter) the editorial decision to publish or reject was dependent on the order in which the referees' reports arrived at a particular journal.

They may be limits to improvements to reliability that can be made by refining peer review process. A study by BMJ showed no evidence that training referees improved the quality of outcome. More generally, a mathematical model by Thurner & Hanel (2010) showed that the presence of even very small fractions of "random" (i.e. incompetent) or "rational" (i.e. biassed towards rejecting papers that might draw attention from their own work) reviewers had a big impact, with reduction in the quality of published papers: "if it can not be guaranteed that the fraction of rational and random referees is confined to a very small number, the peer-review system will not perform much better than by accepting papers by throwing (an unbiased!) coin."

Other researchers have shown low correlation between peer review scores and subsequent citation counts (e.g. Casati et al 2009).

It is worth noting in passing that the use of metrics as alternatives (or supplements) to peer review may be just as unreliable, a point which we shall return to in our discussion of post-publication review.

Peer review has also been shown to be poor at detecting errors. For example, BMJ conducted a study (Godlee et al 1998) in which eight errors were deliberately introduced into a paper which was then reviewed by 220 referees. The average number of errors spotted was two, no-one spotted more than five, and about a sixth failed to spot any.

Many commentators regard traditional peer review, particularly in its single-blind form as offering too much scope for bias. Evidence has been presented for nationality bias, language bias (particularly Anglocentrism) and specialty bias (e.g. see the JAMA special issue 280(3) 1998). There is also strong evidence for publication and outcome bias (Dwan et al 2008) which are of particular concern in medical trials. Many also believe that peer review tends to favour the status quo, maintain consensus, suppress dissent and generally discourage innovation, especially when this come from outside the "insider network" (Cope & Malantziz 2009).

More encouragingly, although there has been tentative evidence for gender bias in the past, more recent studies (Marsh et al 2008; Marsh & Bornmann 2010) have shown gender to have no effect on outcome.

On the other hand, Marsh et al showed that using referees nominated by the researchers created a major systematic bias leading to inflated, unreliable and invalid ratings. This study concerned peer review of grant application rather than journal articles but earlier work by Schroter et al (2006) also concluded that editors should be cautious about relying on the recommendations of author-nominated reviewers, and Bornmann and Daniel (2010) found that "editor-suggested reviewers rated manuscripts between 30% and 42% less favourably than author-suggested reviewers".

As well as not being particularly effective at detecting academic misconduct, peer review is also criticised for the opportunities it offers for misconduct by reviewers. In addition to the (conscious or unconscious) biases already discussed, peer review is open to abuse such as delaying (perhaps by requiring further work) or rejecting a paper in order to allow the reviewer to publish first, or for reviewers to plagiarise the author's ideas.

A major area of criticism of peer review is that it is slow, delaying publication unnecessarily. In a networked world where information can be published instantaneously and the research cycle ever shorter, overall submission-to-acceptance times of months can appear archaic and inefficient. Of course, the largest component of the longer delays will normally be the time spent by authors revising the manuscript rather than time taken by the reviewing process itself (many science and medical journals, for instance, have times to first decision of a few weeks) but it is also held that much of the additional work and revisions required by referees is unnecessary (or that with post-publication review revision could take place after publication).

A somewhat philosophical objection to traditional peer review is made by Cope and Malantziz (2009) in a dense but rewarding essay, in which they observe that what is reviewed (the manuscript) is the new knowledge itself (e.g. the data) but a rhetorical presentation of the work, typically written to persuade the reader of its correctness. We shall return to this point later (see Opportunities and challenges).

Overall, therefore, a strong case is made by critics (e.g. Smith 2009) against traditional single-blind review on the grounds of fairness: it is unreliable and hence partly random, open to bias and misconduct and on top of this, is conducted by a reviewer operating behind a cloak of anonymity. When this is combined with evidence of its ineffectiveness (e.g. at detecting errors) and its inefficiency (slow and expensive), and the lack of an evidence base for much of its practice ("the practice of peer review is based on faith in its effects, rather than on facts"), critics say there is an unanswerable case for change. The question is, how should it be changed, and are the proposed improvements or alternatives better than the current system?

Improving peer review

What is PR for?

It is one thing to criticise yet another to propose alternatives. To start with, it is difficult to decide how to improve the system without agreement on its goals (Wager 2006). Is it to select the best papers to publish in a journal? To minimise (if it can't eliminate) fraud and other misconduct? To improve the quality of papers published? To improve the quality of research? To act as a filter, by rejecting bad work, or by deciding *where* a paper is published rather than *whether* it is published?

Lee & Bero (2006) say that "Peer review is not currently designed to detect deception, nor does it guarantee the validity of research findings. It should, however, identify flaws in the design, presentation, analysis and interpretation of science and provide prompt, detailed, constructive criticism to improve research."

Many (if not most) practising editors and publishers would concur and take a pragmatic line that combines many of these goals, probably focusing on selecting the best papers for their journals and improving published papers while attempting to minimise the disadvantages. More radical innovation arises where it is felt that this provides an inadequate response to the criticisms.

It might be helpful to keep in mind what the academic community (specifically in the form of authors and academic readers, and to some extent research funders) wants and needs from peer review. The interests of editors and publishers and of the wider readership are surely secondary.

It is clear that authors currently value the improvements that peer review makes to their published work. In this regard, "authors want specifics regarding problems ... and, when feasible, concrete suggestions for fixing them and for otherwise improving the presentation" (Nickerson 2005), all delivered in a timely fashion.

Academic readers, as the Tenopir et al (2010) survey reminds us, continue to place importance on the screening function (distinguishing peer-reviewed from non-peer-reviewed material) and continue to see value in the journal brand as a signifier of type and quality of content. Readers can also be assumed to benefit from the improvements to presentation and readability and overall quality of published papers (though a few might argue that the improvements are not worth the cost and time involved).

Making peer review more effective and efficient

A number of more-or-less common-sense approaches have been suggested for improving peer review, including being helpful and reasonable to authors and reviewers; rating reviews and giving feedback to reviewers; checklists or templates for reviewers; training reviewers (the evidence for the effectiveness of this is limited, but 56% of reviewers in a survey said there was a lack of guidance on how to review; and 68% thought formal training would help (Sense About Science 2009)); choosing reviewers based on knowledge, expertise and currency, rather than their eminence; and providing recognition and rewards to reviewers (e.g. publishing a list of reviewers; inviting to an annual reception; awards for best reviewer of the year; certificates they can show to their employer; even token payments).

Kent Anderson makes the point that not all peer review reports are done to the same standards and suggests that peer review could be improved if each published paper was accompanied by a disclosure statement covering factors such as: the number of outside reviewers, the degree of blinding, the number of review cycles needed before publication, the duration of the peer review portion of editorial review, other review elements included (technical reviews, patent reviews, etc.), editorial board review, editorial advisers review, statistical review, safety review, ethics and informed consent review (Anderson 2010).

In a similar vein, Jennings (2006) suggested that greater transparency on the part of journals in terms of peer review metrics could foster competition which could drive up standards. He acknowledged, however, that finding metrics covering the benefits (e.g. value added) as well as the costs (e.g. time spent reviewing) that were reliable, easily digestible, economical and resistant to gaming would be a substantial challenge.

Lee & Bero (2006) also advocated improving transparency (e.g. having a published process; requiring disclosure of all competing interests (reviewers and editors, not just authors); clear sanctions for misconduct or failure to disclose); promoting fairness (e.g. having an appeals process or ombudsman); and facilitating review through specific instructions and evaluation tools for reviewers, and through rewards and recognition for reviewers.

It is not possible to review the state of peer review at this point without considering the impact of the open access journal *PLoS ONE*. Its key innovation was a peer review process that explicitly makes no judgement about the value, significance or potential impact of a submitted article, freeing the reviewers to assess just the scientific rigour, combined with a platform designed to support post-publication review. It has also made effective use of cascade review (below), drawing on papers rejected from other high-profile PLoS journals for a sizeable fraction of its submissions. PLoS argues that it is better for significance and impact to be assessed after publication, using both traditional tools (e.g. citations) and post-publication review (comments and ratings) and other new article-level metrics (e.g. usage). The enormous and rapid success of *PLoS ONE* (well over 60,000 authors have submitted to the journal, which now has a reviewer bank of over 35,000, making it one of the largest journals in the world) has left other publishers scrambling to emulate it (e.g. *BMJ Open*, *Nature Communications*, *AIP Advances*, Hindawi's *Scholarly Research Exchange*, and more planned for launch during 2011 from other publishers).

"Cascade" review is being increasingly adopted as a way addressing inefficiencies in the present iterative process of rejection and resubmission elsewhere, with peer review then starting again. In cascade review, the rejected author is asked if they wish to have their paper resubmitted with its previous reviewer reports attached, thereby reducing the amount of reviewing required by the next journal. The benefit for the author is that it may offer a faster route to publication, particularly if the report is of the type "sound work but of insufficiently broad application or significance for this journal"; on the other hand an author that received what they see as unwarranted criticism or requests for further work may prefer to take their chances anew.

Cascade review can take place within a single publisher's list of journals and this is by far the most common variant. It works particularly well when combined with an option like PLoS ONE that offers a faster review process based on scientific rigour rather than significance or importance; PLoS has reported that nearly a third of PLoS ONE's submissions came via the cascade route and that PLoS ONE was the single most common place for authors rejected by other PLoS journals to resubmit (Patterson 2010).

Cascade review can also operate between independent journals with different publishers, as at the Neuroscience Peer Review Consortium. This was launched in 2008, initially for a trial period of year and has recently (November 2010) been reconfirmed by its members for a further 12 months. Participants report (private communications) modest success and is broadly beneficial, and seem happy with the quality of submissions received via this route; on the other hand the numbers of papers are small (e.g. the fraction of rejected authors from the leading journal in the field that chose this route is well under 10%) and so making a limited impact on overall reviewing loads.

There is also a place for using technology to improve the efficiency of peer review. Adoption of online manuscript tracking systems is now widespread (and universal at leading publishers) but their introduction undoubtedly helped reduce the cost of administering the system and the time taken. (Though there has also been a suggestion that the automated emails sent by these systems are more likely to be trapped by anti-spam filters, thus requiring more invitations to be sent to secure the required number of reviewers (Vines et al 2010).)

More recently technology has been introduced to improve the detection of image manipulation; to help identify possible plagiarism (CrossCheck); and to automate the process of referee selection (Collexis).

Making peer review fairer

Issues of fairness, including scope for bias and the lack of transparency and accountability feature highly in authors' complaints about peer review.

One unusually public statement of this was an open letter from 14 stem cell researchers addressed to editors of journals in their field calling on them to "greatly improve transparency, fairness and accountability ... [by publishing] the reviews, response to reviews and associated editorial correspondence ... while preserving anonymity of the referees" (Smith et al 2009).

Many others have suggested reviewers should be made more accountable for their recommendations; there is, for instance, no cost to the reviewer in recommending rejection (or unnecessary work), or simply in doing a poor job, but the consequences for the author can be substantial. The most commonly suggested remedy for this is to adopt some form of open peer review, which we discuss below, or to adopt double-blind review. Other suggestions to increase accountability include an appeals process (Lee & Bero 2006) or right to challenge the review (Franklin 2010); a statement disclosing the scope and extent of the review (Anderson 2010); and even creating metrics to measure reviewing performance (Aspinall 2009).

Open peer review

As we saw above, researchers when surveyed appear to support quite strongly the idea of double-blind review, most likely because because it was perceived to be fairer, and double-blind review is already popular in some disciplines. It does not seem, however, to be receiving the attention that open peer review is getting as an alternative to single-blind review, perhaps because of the practical difficulties of implementation.

It is open peer review that is currently most often advocated to address the issues of transparency, fairness and accountability.

The term open review is used confusingly to refer to several related but different variants of peer review, and it is worth taking a moment to identify these since it is not always clear in debates over the merits of open peer review exactly what is being referred to. (See also Hodgkinson 2007.) The "openness" in review can refer to:

- the lack of blinding of the reviewer's identity. There can be several variants of this: the reviewers' names can be made known to the author (that is, the reports are signed) but are not made public (as at the BMJ, the *Journal of the Royal Society of Medicine* and BioMed Central medical journals), or alternatively the reviewers' names can be made public and attached to the published paper. It is also possible for the reviewers to remain anonymous during the process of review but for their names subsequently to be revealed and published alongside the paper. Another variant is to allow the author to nominate the referees (as at *Biology Direct*).
- access to the reviewer's reports and the associated author responses and other documentation. The review process can take place in the open, with the submitted manuscript, reviewer reports, author responses and editor's comments etc. made available in real time; this is the approach adopted at the journal *Atmospheric Chemistry & Physics*. Alternatively, the reviewing documents can be published simultaneously with the paper (e.g. as at *The EMBO Journal*). An additional

requirement can be that the reviewer does not provide confidential comments to the editor in addition to the report that is made available to the author.

- the pool of people able to comment on the submitted manuscripts, i.e. opening this up to the wider community rather than relying on two or three reviewers selected by the editor. This approach can take place after publication (see Post-publication peer review, below) or prior to publication (as at *Atmospheric Chemistry & Physics* or at the *Nature* peer review trial). The intended reviewer pool can be the wider research community (perhaps including regions under-represented among reviewers), statisticians or other professionals, or even patients as proposed at the *Journal of Participatory Medicine* (Shashok 2010).

The advantages and disadvantages of open peer review have been widely studied and debated in the literature (e.g. see Groves 2010; Khan 2010). The main arguments in favour generally start from the clear evidence of bias arising from the lack of transparency in single-blind review, as discussed above. This could be addressed by double-blind review but studies (e.g. Godlee et al 1998; van Rooyen et al 2010) have shown that there is no advantage of double-blind over traditional or open review and that there are practical difficulties in administering it (e.g. reviewers can identify the author in 10-20% of cases from internal clues). Supporters of open review also believe that it is fundamentally unfair ("kafkaesque") for important decisions potentially affecting the author's career to be made behind a veil of anonymity. Supporters also argue that open review will increase accountability which in turn will improve the quality of reviews, as reviewers will take more care and avoid superficial work when their review is signed. It would also give more recognition to reviewers for their work.

One key argument against open review is that it is not realistic to expect junior reviewers to criticise eminent authors (who may determine the reviewer's future career) as strongly. Even where there is not a power asymmetry, reviewers may pull their punches for fear of giving offence. There is some evidence that reviewers are slightly more likely to give positive recommendations in signed than anonymous reviews. It is also clear that many reviewers are uneasy about open review and at journals where they are encouraged to sign reviews but allowed to remain anonymous if they prefer, many do exercise this choice.

One reason for taking care to distinguish the variants in open peer review is that the degree of support for each varies substantially among the academic community. In particular, reviewers remain reluctant to have their names publicly disclosed and attached to the published article. For example, about half researchers responding to our survey said this would make them less likely to review for a journal (Ware 2008), and van Rooyen et al (2010) found in more controlled conditions that a similar proportion of potential reviewers would refuse if their name was to be made public. In contrast, Richard Smith reported that when the BMJ adopted its current open review system where reviewers sign their reviews but do not have their names disclosed publicly, only a handful of the 5000-odd reviewers on their database declined (Smith 2009). It does seem therefore that reviewers may be prepared to have their names disclosed to the author but are mostly wary of publicly signing their reviews.

Willingness to sign reviews appears to vary by discipline too. Bloom (2006) reported that there was greater acceptance of open signed review in the Biomed Central medical journals than in their biological science journals, while at *Atmospheric Chemistry & Physics* theoreticians and modellers (40%) were far more likely to sign reviews than experimentalists (10%).

Nature's trial of open peer review (Greaves et al 2006) is frequently cited as evidence of reluctance by the scientific community to adopt open review. Certainly the particular kind of open review offered (and specifically labelled as a trial) was not much valued or used, but this finding may be of limited application. The trial invited authors to allow their submitted manuscripts to be posted for open comments (which had to be signed) while receiving conventional closed review in parallel. Only 5% of authors agreed to participate. Although web traffic to the review pages was significant, there were typically very few comments and most of these were not technically substantive.

The successful implementation of different kinds of open peer review at the BMJ, the Biomed Central medical journals and the various examples reviewed below suggests that the choice of type of open review system is important.

One example of a (large) journal that has successfully implemented open peer review in the form of publishing the reviewer reports and associated peer review materials ("process files") alongside the article but maintaining reviewer anonymity is *The EMBO Journal*. Its Chief Editor Bernd Pulverer reported (2010) that the policy was launched in January 2009 with the expected benefits of showcasing the quality and thoughtfulness of reviews, acting as a teaching tool, and encouraging reviewers to take more care. It appears the changes were accepted by the community, with submissions steady and the acceptance rate about the same (~10%) and only a small fraction (~5%) of authors declining to participate. Several reviewers said they took more care. The materials are used: almost all had been viewed, with process files being viewed about a tenth as often as the main paper. The main disadvantage was the additional time taken to process the materials, about 1.5 hours per articles. Pulverer says that EMBO considered double-blind review but could not see how to implement without adding delays or additional work for authors. They also considered having reviewer reports signed (though not publicly) but felt that "the stakes are too high" in the competitive world of biological research.

The "Frontiers" journals attempt to balance the interests of anonymity and openness by keeping reviewers anonymous during the review process but revealing their names (which are published with the article) on publication.

Atmospheric Chemistry & Physics (owned by the European Geophysical Union and published by Copernicus) has successfully pioneered since 2001 an unusually open peer review system (Pöschl 2010; Koop & Pöschl 2006). Two stages of review are employed. In the first stage, submitted manuscripts are first screened for general suitability and the manuscript is published by the journal as a "discussion paper". Discussion remains open for 8 weeks, during which time the (editor-commissioned) reviewers' reports, author responses and community/public comments are all made public. Reviewer reports and comments can be anonymous or signed; about three-quarters of reviewers choose not to disclose their names. In the second stage, manuscript revisions and follow-up review proceeds privately as in traditional review. If accepted the paper is published in the main journal. If not accepted, the discussion paper and associated comments are permanently archived.

Each paper gets 4–5 interactive comments, with comments in total equalling about 50% of the extent of the paper. Rejection rates are lower than journals of comparable scientific quality, possibly because the open review deters submission of carelessly prepared manuscripts. "Only a handful" of the 10,000 comments received had to be removed or replaced because of inappropriate wording.

EGU regards the system as very successful and has extended it to a dozen or so sister journals. Pöschl describes the advantages as follows:

- using open review allows the discussion paper to be treated as the formal publication date; the time spent on review and revision does not, therefore, delay publication giving rapid dissemination of results
- direct feedback and public recognition of good papers
- prevention of misconduct by reviewers such as hidden obstruction or plagiarism
- citable documentation of critical comments, controversial arguments, scientific flaws and complementary information.

Boldt (2010) has described using the arXiv repository to create a similar system for mathematical and physical sciences though at this stage this remains a hypothetical proposal.

The journal *Biology Direct* publishes signed reviewer reports plus author responses (Koonin 2006). The peer review process is also unusual in that authors choose the reviewers from the editorial board. The paper is published if the author is able to obtain three such reviews (even if the reviews are critical or negative), otherwise it is rejected. (Authors can also "self-reject", that is withdraw their paper if they do not to proceed with negative reviews attached to their paper.)

The *Journal of Interactive Media in Education* (http://jime.open.ac.uk) also uses a two-stage open review process. In the first stage, reviewers post reviews to a private discussion site open to just authors and reviewers. Reviewers may remain anonymous but generally sign their reviews. Authors and reviewers are encouraged to discuss the points raised. In the second stage, if the editors believe the first-stage discussion warrants it, the article is published as a preprint for public open peer review, with the author–reviewer discussion forming the seed of this discussion, which last for a month. The editor then posts a report summarising the main issues and specifying any change requirements to the authors.

Post-publication peer review

Post-publication peer review is a somewhat slippery concept. Where, for instance, is the hard-and-fast distinction between a formal Comment, a Rapid Response, a moderated comment on the journal website and a response published on on a separate blog but visible on the journal website via the "trackback" protocol? Much of what is sometimes referred to as post-publication review could also be considered as part of the normal academic discourse brought forward into the internet age (e.g. see the *Nature* editorial on the NASA arsenic microbe affair (Nature 2010)). On the other hand, the system employed by *Atmospheric Chemistry & Physics* could also be regarded as post-publication review, since it defines publication as the date at which the "discussion paper" (i.e. submitted manuscript) is posted.

For this article, we shall treat post-publication review as the systems (actual and proposed) to review or critique and filter journal articles using comments or similar, ratings and other article-level metrics including usage and citation data.

There are a number of philosophies and motivations (or if you will, "memes") behind the arguments for post-publication review. Ideas drawn from the Web 2.0 world figure strongly, particularly the "wisdom of the crowds" and for those who see the possibility of replacing (rather than just supplementing) conventional review, the Clay Shirky soundbite "publish then filter" has resonance.

More specifically, it is argued that allowing readers to become reviewers by annotating and commenting on papers and by rating them will broaden the range of viewpoints and insight ("why have just two reviewers when you can have many?"). Conventional peer review is held to be unnecessarily slow: rapid publication without elaborate reviewing and lengthy rework/revision followed by community critique would accelerate the dissemination of research findings.

There are arguments of efficiency in favour of the "publish then filter" model. It is pointed out that a significant proportion of the literature is never cited; reviewing this work (other than perhaps a light pre-screening) would thus appear a waste of resources, particularly as peer review has been demonstrated to be highly unreliable (Neylon 2009; Smith 2009).

Another motivation for expanding the range of available article-level metrics has been to escape the dead hand of the (misused) Impact Factor.

The problem with commenting systems is that in most journals where they have been adopted, researchers have been very reluctant to make use of them. For example, *Nature* strongly encourages commenting and post-publication discussion and yet reports that "of the thousands of papers published every year, only a few attract substantive comments" (Nature 2010). PLoS is similarly if not more keen on promoting commenting and rating of articles, and yet the proportion of its published articles with any rating at the time of writing was about 5% and the proportion with any comment was 11% (my analysis from PLoS data). Of equal concern is a study finding that authors rarely bothered to respond to online criticisms of their work (Gøtzsche et al 2010).

From most researchers' perspective there seems very little incentive to comment; there is a lack of time in which to do it; and there is limited value in comments placed by others. For a researcher that currently reviews 1-2 papers and reads 20-30 papers a month, a move to post-publication review only would mean having to read the 20-30 papers with the same degree of critical attention currently given to the 1-2 that were reviewed, a substantial increase in workload.

More generally, several studies (e.g. RIN 2010) have found researchers to have made low take-up of Web 2.0 applications, with only a small fraction (13% in the RIN report) frequently using tools such as blogs, wikis and so on. The RIN report found a major set of barriers revolved around perceptions of quality and trust. Overall its authors concluded that there was "little evidence at present to suggest that web 2.0 will prompt in the short or medium term the kinds of radical changes in scholarly communications advocated by the open research community". Nonetheless, the report also found that "a significant minority of researchers believed that peer review in its current forms would become increasingly unsustainable over the next five years, and nearly half (47%) expected that it would be complemented by citation and usage statistics, and user ratings and comments".

There may be more fundamental reasons for caution in the use of metrics as a way of filtering articles.

To start with, it has been shown that the ubiquitous 5-star rating tool used across the web and on some journal platforms is in practice flawed and produces highly skewed results. The vast majority of ratings on YouTube, for instance, are 5 stars, and it turns out this is repeated across the web with an overall average estimated at about 4.3 stars (Crotty 2009). Analysing the data for articles published by PLoS, for example, gives an average (for the 5% of articles that have been rated) precisely in line with this wider experience.

The problem is that rating systems select for popularity rather than quality, which is almost the antithesis of what is normally wanted for scholarly evaluation. (For example, the most-rated PLoS article by some margin (with an average rating of 4.9, no less) is an essay on how the "granting system turns young scientists into bureaucrats and then betrays them".) There is an obvious positive feedback loop, amplifying ratings and reducing diversity in a "winner takes all" system.

Another problem with commenting and rating systems is that they are open to gaming and manipulation. For instance, the Amazon system has been widely abused and it has been demonstrated how easy it is for an individual and small group of friends to influence the popularity metrics even on hugely-trafficked websites like the *New York Times*.

Despite the drawbacks, there does seem to be a place for a wider range of metrics than just the journal Impact Factor, almost certainly as a supplement to traditional peer review rather than a replacement. The informal Alt-metrics group (Priem et al 2010) proposes the development of new metrics and tools for their adoption.

It does seem important that if metrics are used, they should be many and diverse. Nielson (2010) argues that it is always better to use a homogeneous range than any single metric, no matter how good it may be. Centralised metrics, he argues, suppress diversity, create perverse incentives and misallocate resources. It may be, of course, that if you wait for more considered metrics (such as citations and usage patterns) to accumulate to balance more populist ratings and comments that there is no time advantage compared to the delays incurred by peer review.

One criticism of (proposed) post-publication systems that rely on metrics in place of conventional peer review is that they do not provide the interactive process for improving papers prior to publication which as we saw above is seen by researchers as a key benefit of peer review (e.g. Harnad 2009). The existing system of stratified journals provides a framework within which peer review can be conducted to different quality levels.

One proposed way of improving rating and commenting systems is to weight such ratings according to the reputation of the rater (for example as done on Amazon, eBay, Wikipedia or Slashdot etc.). Frishauf (2009) proposed a reputation system for peer review in which the review (whether pre- or post-publication) would be undertaken by people of known reputation (or trustworthiness, to be embodied in a "trustmark"). In general, reputation systems are intended to achieve three things: to foster good behaviour, punish bad behaviour, and reduce the risk of harm to others cause by bad behaviour (Ubois 2003). Key features are that only individuals have reputations; reputations can rise and fall (although as Mark Twain observed: "once you have a reputation for being an early riser, you can sleep in to noon everyday"); reputations is based on behaviour rather than social connections (which distinguishes reputation from web-of-trust systems). Reputation systems do not have to operate using the true names of participants but effective reputation systems must be tied to a robust and enduring identity infrastructure (otherwise an individual could simply walk away from a poor reputation). At this stage it seems unlikely that the research community is ready to embrace such systems *en masse*.

One problem with reputation systems is that having a single formula to derive the reputation leaves the system open to gaming. Gasher (2008) proposed a decentralised system in which reviews (which could be independent of publication) would be securely digitally signed by the reviewer, with the digital signature linking the review and paper. A decentralised web of reviewers and papers is thus created which could be data-mined to reveal information about the influence or well-connectedness of individual researchers within the research community. Depending on how the data were mined, this could be used as a reputation system or web-of-trust system or some combination, and would be resistant to gaming because it specified no particular metric. The idea remains a hypothetical one at this stage, however.

Philica, a small online journal started by two UK psychologists allows articles to be posted on any subject without any prior screening and then encourages its users to review them. It uses a reputation system to weight the reviews (including weighting by average scores of the reviewers' own publications; reviewer status (academic affiliations score higher) and the age of the review (new reviews count higher). The results are not especially encouraging, with the journal (including its Most Popular list) riddled with pseudoscience.

A wider range of metrics at the article rather than journal level is, however, almost certainly a good thing for peer review. The focus on Impact Factor creates problems for peer review by supporting what Sir John Sulston called an "inefficient treadmill" of resubmissions to the journal hierarchy. This increases the amount of reviewing and delays the communication (or at least the formal publication) of important results.

As Harnad (2009) points out, though, this is not a zero-sum game: getting more value from post-publication metrics does not mean having to throw out the value the community gets from pre-publication peer review.

Opportunities and challenges

One immediate challenge for peer review is dealing with research data. We noted above that a paper is a rhetorical (re)presentation of the academic work described, at one or more steps removed from the work itself. In a digital networked world this need not (in principle) be the case; for example, complete datasets can be made available alongside their interpretation. In our survey (Ware 2008) we found that a majority (about two-thirds) of researchers and journal editors agreed that review of associated data was desirable in principle, though some 40% thought it was unrealistic to ask reviewers to take on this challenge. Journals are increasingly allowing (or encouraging, or in some cases requiring) data to be deposited as supplementary materials or linked to via external repositories. There are some areas where journals already effectively review data, for example some of the International Union of Crystallography journals or the *Signaling Gateway* database.

Rather going against this trend, or perhaps offering a warning of the challenges that may lie ahead for data review, the *Journal of Neuroscience* this year decided to stop allowing authors to include supplemental material because they felt it was undermining peer review (Maunsell 2010). The journal argued that the sheer volume of material and its rate of increase made it increasing impractical for reviewers to deal with; in practice, therefore they would either review it superficially (or not at all), leading to un-peer-reviewed material being published under the journal's name, or that time would be taken from the review of the main paper, to its detriment. It encouraged reviewers to demand "subordinate or tangential" additional work from authors, thus causing unnecessary work and delay. Lastly it also undermined the concept of a self-contained research report (though some might feel this begs the question of whether this is a desirable in a networked world).

Phillip Bourne, Editor-in-Chief of *PLoS Computational Biology* goes further. In "What Do I Want from the Publisher of the Future?" (Bourne 2010) he argues that journals should publish the entire scientific workflow (i.e. idea, hypothesis, experiment, data, analysis, results, discussion, conclusion, publication) rather than just the final paper. He acknowledges that this would create challenges for peer review but believes that with appropriate tools it would be feasible.

Another challenge arises from the rapid growth in output of research papers from East Asia, particularly China. Most journals are based in the West and have likely relied historically on a reviewer base reflecting the previous US/European domination of research output. Clearly that has to change but with a relatively young researcher population that lacks extensive experience of peer review and that may face challenges in reviewing in English, there is likely to be a lag before these new research populations fully come on stream as reviewers.

One suggestion for improving peer review that has been more written about than attempted is to use market mechanisms. The main motivations are either to create greater reviewer accountability (by requiring them to back their judgements with some kind of virtual currency or similar (Robinson 2009; Neylon 2010), or avoid a "tragedy of the commons" in which authors take advantage of peer review but do not contribute, by imposing submission fees (normally in virtual currency) which are earned by reviewing. Some bepress journals do operate the latter approach but the pan-journal system envisaged by Fox & Petchey (2010) with a centralised bank to manage the virtual currency seems little more than speculation.

Another notion discussed for some time but yet to find much substance is that of the disaggregated journal. In this model, the publishing functions are separated, with *registration* and *dissemination* provided by a repository (such as the arXiv) and *certification* (essentially a combination of peer review and branding) conducted by third parties, thus creating so-called overlay journals (e.g. Van de Sompel 2006). It has also been suggested that learned societies might offer peer review/certification services as a future core function, possibly on fee-per-article basis. It seems just as plausible, however, that publishers might seek to build communities of authors and reviewers around journals as they compete more for their attention and their "hearts and minds".

Let us return now to the question we opened with: is peer review in crisis? As we have seen, there are three main arguments given to support this claim: that it does not work (unreliability, bias, etc.); that it is too slow, delaying the progress of research; and that reviewers are over-burdened to the point where they can no longer cope, especially as outputs continue to rise.

There seems to be very little data to support the last point at the macro level (individual journals may have their own problems, of course, and peer review of grant applications may be a different case). Large-scale surveys show that the average frequency of reviewing is less than once per month, with more active reviewers perhaps reviewing a couple of papers a month on average. This is a real professional commitment but does not on the face of it appear to be an unsupportable burden. There is also no reason in principle why the burden should increase as the the number of papers published increases, assuming this follows the historical pattern, because the number of potential reviewers increases at the same rate as the number of papers (or possibly even faster, if the number of co-authors per paper continues to rise). The journal *Molecular Ecology*, for instance, reported that its pool of reviewers rose in line with submissions; within its own community each paper had an average 4.5 authors and required 2.7 reviews, so only 0.6 reviews per author was required to support each new paper.

It is true that peer review takes time, and the inefficient process of multiple submissions down the journal hierarchy encourage by the misuse of journal metrics and research assessment is a notable waste of that most precious resource.

More generally it is also true that it is easy to enumerate the defects of the current system but not so easy to find anything to replace it. As Richard Smith says:

> *"... peer review is a flawed process, full of easily identified defects with little evidence that it works. Nevertheless, it is likely to remain central to science and journals because there is no obvious alternative, and scientists and editors have a continuing belief in peer review."*

Instead of a collapse, therefore, we see a wide range of debate, experimentation and innovation across a growing number of journals. Far from being in crisis, it could be argued that peer review has never looked more vibrant in its growing diversity.

References

Anderson, K (2010) Improving Peer Review: Let's Provide an Ingredients List for Our Readers, Scholarly Kitchen blog 30 March 2010 http://scholarlykitchen.sspnet.org/2010/03/30/improving-peer-review-lets-provide-an-ingredients-list-for-our-readers/

Armbruster, C (2007) Moving out of Oldenburg's Long Shadow: What is the future for society publishing?, Learned Publishing 20(4), 259-266 http://dx.doi.org/10.1087/095315107X239627

Aspinall, W (2009) A metric for measuring peer-review performance, Peer-to-peer blog http://feeds.nature.com/~r/peer_review/rss/peer_to_peer/~3/fIAYO5-wPVg/a_metric_for_measuring_peerrev.html

Björk, B-C et al (2009) Scientific journal publishing: yearly volume and open access availability, Information Research 14(1) http://informationr.net/ir/14-1/paper391.html

Bloom, T (2006) Online frontiers of the peer-reviewed literature, Nature Peer Review Debate http://dx.doi.org/10.1038/nature05030

Boldt, A (2010) Extending ArXiv.org to Achieve Open Peer Review and Publishing, arXiv:1011.6590v1 http://arxiv.org/abs/1011.6590

Bornmann, L & Daniel, H-D (2009) The luck of the referee draw: the effect of exchanging reviews, Learned Publishing 22, 117–125 http://dx.doi.org/10.1087/2009207

Bornmann, L & Daniel, H-D (2010) Reliability of reviewers' rating when using public peer review: a case study, Learned Publishing 23, 124–131 http://dx.doi.org/10.1087/20100207

Bornmann, L & Daniel, HD (2010) Do author-suggested reviewers rate submissions more favorably than editor-suggested reviewers? A study on Atmospheric Chemistry and Physics, PLoS One 5(10):e13345 http://dx.doi.org/10.1371/journal.pone.0013345

Bornmann, L et al (2010) A Reliability-Generalization Study of Journal Peer Reviews: A Multilevel Meta-Analysis of Inter-Rater Reliability and Its Determinants, PLoS ONE 5(12): e14331 http://dx.doi.org/10.1371/journal.pone.0014331

Bourne, P (2010) What Do I Want from the Publisher of the Future?, PLoS Comput Biol 6 (5): e1000787 http://dx.doi.org/10.1371/journal.pcbi.1000787

Casati, F et al (2009) Is peer review any good? A quantitative analysis of peer review, University of Trento Technical Report #DISI-09-045 http://eprints.biblio.unitn.it/archive/00001654/01/techRep045.pdf

Cope, B & Malantziz, M (2009) Signs of epistemic disruption: Transformations in the knowledge system of the academic journal, First Monday 14(4) http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2309/2163

Crotty, D (2009) How Meaningful Are User Ratings? (This Article = 4.5 Stars!), http://scholarlykitchen.sspnet.org/2009/11/16/how-meaningful-are-user-ratings-this-article-4-5-stars/

Dwan, K et al (2008) Systematic review of the empirical evidence of study publication bias and outcome reporting bias, PLoS ONE 3(8): e3081 http://dx.doi.org/10.1371/journal.pone.0003081

Fox, J & Petchey, O (2010) Pubcreds—Fixing the Peer Review Process by 'Privatizing' the Reviewer Commons, Bulletin of the Ecological Society of America 91:325–333 http://dx.doi.org/10.1890/0012-9623-91.3.325

Franklin, R (2010) Putting peer review on trial, Physics World 23(12) 17

Frishauf P (2009) Reputation Systems: A New Vision for Publishing and Peer Review, J Participat Med. 2009(Oct);1(1):e13a http://www.jopm.org/opinion/commentary/2009/10/21/reputation-systems-a-new-vision-for-publishing-and-peer-review/print/

Gasher, M (2008) GPeerReview - a tool for making digital-signatures using data mining, KDnuggets http://www.kdnuggets.com/news/2008/n16/5i.html

Godlee, F et al (1998) Effect on the Quality of Peer Review of Blinding Reviewers and Asking Them to Sign Their Reports, JAMA 280:237-240 http://dx.doi.org/10.1001/jama.280.3.237

Greaves, S et al (2006) Nature's trial of open peer review, Nature Online Peer Review Debate http://dx.doi.org/10.1038/nature05535

Groves, T (2006) How can we get the best out of peer review?, Nature online debate on peer review http://dx.doi.org/10.1038/nature04995

Groves, T (2010) Is open peer review the fairest system? Yes, BMJ 2010; 341:c6424 http://dx.doi.org/10.1136/bmj.c6424

Gøtzsche, PC et al (2010) Adequacy of authors' replies to criticism raised in electronic letters to the editor: cohort study, BMJ 2010; 341:c3926 http://dx.doi.org/10.1136/bmj.c3926

Harnad, S (2009) Post-Publication Metrics Versus Pre-Publication Peer Review, Author's blog http://openaccess.eprints.org/index.php?/archives/612-Post-Publication-Metrics-Versus-Pre-Publication-Peer-Review.html

Hodgkinson, M (2007) Open peer review & community peer review, Author's blog posting http://journalology.blogspot.com/2007/06/open-peer-review-community-peer-review.html

Jefferson, T et al (2007) Editorial peer review for improving the quality of reports of biomedical studies, Cochrane Database Syst Rev. 2007 Apr 18;(2):MR000016 http://www.ncbi.nlm.nih.gov/pubmed/17443635

Jennings, C (2006) The true purpose of peer review, Nature Online Peer Review Debate http://dx.doi.org/10.1038/nature05032

Journal of Interactive Media in Education, Peer Review Policy, http://jime.open.ac.uk/jime/about/editorialPolicies#peerReviewProcess

Khan, K (2010) Is open peer review the fairest system? No, BMJ 2010;341:c6425 http://dx.doi.org/10.1136/bmj.c6425

Koonin, E et al (2006) Reviving a culture of scientific debate, Nature Online Peer Review Debate http://dx.doi.org/10.1038/nature05005

Koop, T & Pöschl, U (2006) An open, two-stage peer-review journal., Nature Online Peer Review Debate doi:10.1038/nature04988

Lee, K & Bero, L (2006) Increasing accountability, Nature Online Peer Review Debate http://dx.doi.org/10.1038/nature05007

Marsh, H & Bornmann, L (2009) Do women have less success in peer review?, Nature 459, 602 http://dx.doi.org/10.1038/nj7246-602a

Marsh, HW et al (2008) Improving the Peer-Review Process for Grant Applications, American Psychologist, Vol 63(3), 160-168 http://dx.doi.org/10.1037/0003-066X.63.3.160

Maunsell, J (2010) Journal of Neuroscience Announcement Regarding Supplemental Material, J. Neurosci. 30(32):10599 –10600 http://www.jneurosci.org/cgi/reprint/30/32/10599

Nature (2010) Response required, Editorial, Nature 468, 867 http://dx.doi.org/10.1038/468867a

Neylon, C (2009) What Is The Cost Of Peer Review? Can We Afford (Not To Have) High Impact Journals?, Author's blog http://blog.openwetware.org/scienceintheopen/2009/03/03/what-is-the-cost-of-peer-review-can-we-afford-not-to-have-high-impact-journals/

Neylon, C (2010) Peer review: What is it good for?, Author's blog http://cameronneylon.net/blog/peer-review-what-is-it-good-for/

Nickerson, RS (2005) What Authors Want From Journal Reviewers and Editors, American Psychologist, Sept 2005, 661-662 http://dx.doi.org/10.1037/0003-066X.60.6.661

Nielsen, M (2010) The mismeasurement of science, Author's blog posting http://michaelnielsen.org/blog/the-mismeasurement-of-science/

Patterson, M (2009) PLoS Author Research, Private communication

Priem, J et al (2010) Alt-metrics: a manifesto, Website http://altmetrics.org/manifesto/

Pulverer, B (2010) Transparency showcases strength of peer review, Nature 468, 29-31 http://dx.doi.org/10.1038/468029a

Pöschl, U (2010) Interactive Open Access Publishing and Peer Review: The Effectiveness and Perspectives of Transparency and Self-Regulation in Scientific Communication and Evaluation, Liber Quarterly 19, 3/4, 293–314

RIN (2010) If you build it, will they come? How researchers perceive and use web 2.0, Research Information Network report http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/use-and-relevance-web-20-researchers

RIN/CEPA (2008) Activities, costs and funding flows in the scholarly communications system in the UK, Research Information Network report http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/activities-costs-and-funding-flows-scholarly-commu

Robinson, R (2009) On the Design and Implementation of a Market Mechanism for Peer Review and Publishing, The University of Queensland School of Information Technology and Electrical Engineering, Technical Report Number: QRL-2143 http://nicta.com.au/people/rrobinson/publications/citemine-paper.html

Rothwell, PM & Martyn, CN (2000) Reproducibility of peer review in clinical neuroscience, *Brain* 123: 1964–1969 http://dx.doi.org/10.1093/brain/123.9.1964

Schroter, S et al (2006) Differences in review quality and recommendations for publication between peer reviewers suggested by authors or by editors, Journal of the American Medical Association, 295, 314–317

Sense About Science (2009) Peer Review Survey 2009: Preliminary Findings, http://www.senseaboutscience.org.uk/index.php/site/project/395

Shashok, K (2010) Who's a Peer? Improving Peer Review by Including Additional Sources of Expertise, J Participatory Medicine, 2 (Dec 2010) http://www.jopm.org/opinion/commentary/2010/12/08/1278/print/

Smith, A et al (2009) Open letter to Senior Editors of peer-review journals publishing in the field of stem cell biology, EuroStemCell project website http://eurostemcell.org/commentanalysis/peer-review

Smith, R (2006) Peer review: a flawed process at the heart of science and journals, J R Soc Med. 2006 April; 99(4): 178–182 http://dx.doi.org/10.1258/jrsm.99.4.178

Smith, R (2009) In Search of an Optimal Peer Review System, J Participat Med. 2009(Oct); 1(1):e13 http://www.jopm.org/opinion/2009/10/21/in-search-of-an-optimal-peer-review-system/

Tenopir, C et al (2010) Research Publication Characteristics and Their Relative Values, Report for Publishing Research Consortium http://www.publishingresearch.net/projects.htm

Thurner, S & Hanel, R (2010) Peer-review in a world with rational scientists: Toward selection of the average, arXiv:1008.4324v1 http://arxiv.org/abs/1008.4324

Ubois, J (2003) Online Reputation Systems, Release 1.0 21(9) http://cdn.oreilly.com/radar/r1/10-03.pdf

Van de Sompel, H (2006) Certification in a digital era - What functions do we take for granted in print?, Nature Online Peer Review Debate http://dx.doi.org/10.1038/nature05008

van Rooyen, S et al (2010) Effect on peer review of telling reviewers that their signed reviews might be posted on the web: randomised controlled trial, BMJ 2010; 341:c5729 http://dx.doi.org/10.1136/bmj.c5729

Vines, T et al (2010) No crisis in supply of peer reviewers, Nature 468, 1041 http://dx.doi.org/10.1038/4681041a

Wager, E (2006) What is it for? Analysing the purpose of peer review, Nature Online Peer Review Debate http://dx.doi.org/10.1038/nature04990

Ware (2008) Peer review in scholarly journals: Perspective of the scholarly community – an international study, PRC Report, Publishing Research Consortium http://www.publishingresearch.net/PeerReview.htm

Ware, M & Mabe M (2009) The STM report: An overview of scientific and scholarly journals publishing, Report for the International STM Association www.stm-assoc.org/2009_10_13_MWC_STM_Report.pdf

Ware, M (2009) Current peer review practice and perceptions: the view from the field, http://www.against-the-grain.com/2009/08/toc-v-213/