

ICSTI Public Conference 2007 – Session 1

Quality, certification and peer review

Perceptions and misperceptions: Attitudes to peer review

Adrian Mulligan
Elsevier

Assessing the quality of research is of critical importance. In this article I will discuss various research projects conducted at Elsevier, which help us better understand the attitudes towards peer review and how we might improve this most critical of mechanism for assaying research quality.

Background

All the research projects discussed here were conducted in collaboration with either one of two independent research agencies based in the UK, GfKNOP or Illuminas. This discussion will draw upon three key studies. A qualitative study composed of focus groups and teledepth interviews completed in 2003–2004 which delved into the motivations and perceptions towards peer review specifically the opinions of reviewers and editors. A second quantitative study was an online survey which was completed in 2005, and based upon over 6,000 responses (authors, researchers, editors). Thirdly, an online study of Elsevier's own reviewers which commenced in 2007, and thus far has had over 30,000 responses. All the studies are global, and cut across a range of subjects and thus are representative of research community at large.

A brief history and definition of peer review

Peer review has a long history, and has been a method of evaluation since Greek times and is evident in Aristotle's principle of Endoxos [1]. There is almost a two-thousand year hiatus before the first scholarly journals emerge and with it a formalised process of evaluation. The first issue of *Philosophical Transactions* appeared in 1665, and it is widely accredited as being the first journal to formalise peer review ([5], pp. 68–69). However, peer review did not become a standard in scholarly communication until after World War II and most agree with John Ziman ([4], p. 111) that it is “*the lynchpin about which the whole business of Science is pivoted*”.

Despite the wide acceptance of peer review there is no single authoritative definition of peer review, at its simplest it is most inclusive: Carin Olson defines peer review as “*the assessment by an expert of material submitted for publication*” ([3], p. 356–358).

Purpose of peer review

According to the reviewers interviewed in our qualitative study in 2005 there are five key elements to peer review:

- It should assess the novelty of the science;
- Ensure that a consistent and appropriate methodology is used;
- Recent reputable work in the area is correctly acknowledged and referenced;
- Prevent authors from making outlandish claims on minimal results;
- Where feasible, identify incidences of plagiarism.

Though not specifically stated by reviewers, but implicit was that review would help improve the article and identify good quality research. It was also clear what peer review should not do – detect fraudulent submissions. When asked why, respondents said that to detect fraud they must rely on the raw data which hardly is ever available. They agreed that fraud can only be detected when another investigator attempts to duplicate the work. Reviewers also agreed that it was not their role to improve the English of an article.

Motivations for reviewing

To ascertain the effectiveness of peer review it is instructive to understand the motivations driving reviewers. Our research from 2004 showed us that the key motivation is interest, they wish to know what is going on in their field, they wish to keep up date with the latest research and importantly know what the competition is up to. But mostly, reviewers feel duty bound to review as active researchers themselves. It is part of a social understanding; namely that if you believe there should be reviews then you will need to do them. They recognize that publishing their own research will require someone to review their work, thus they must reciprocate in order for their work to be reviewed. A small number of reviewers believed that reviewing on a journal would mean that their own submitted manuscripts would be considered more favourably. For young reviewers being invited to review is a confirmation of standing – acceptance by the field. Being selected by a prestigious journal is quite an accolade.

There are some marginal factors which affect a referee's desire to be involved in the reviewing process; such as building up relations with editors, having their name associated with prestigious journals, being able to demonstrate they have a wide interest in their subject area an important consideration in career development.

There are a variety of motivations for reviewing, but it is clear that the amount of time given to reviewing is closely related to the ebbs and flows of a researchers' career. The younger they are, the more time they have to review articles. Time pressure increases as a career develops and the most expert have less time to review, but time scarcity drops dramatically at the end of a researcher's career, when they reach retirement age and can give more of their time to reviewing. In Elsevier's experience, the best reviewers tend to be quite young or near or past retirement age (see Fig. 1).

Peer review is important and supported

Results from our study in 2005 show that the community at large believe that there is a need for refereed journals; the majority consider it important and are ready to commit time to the review process,

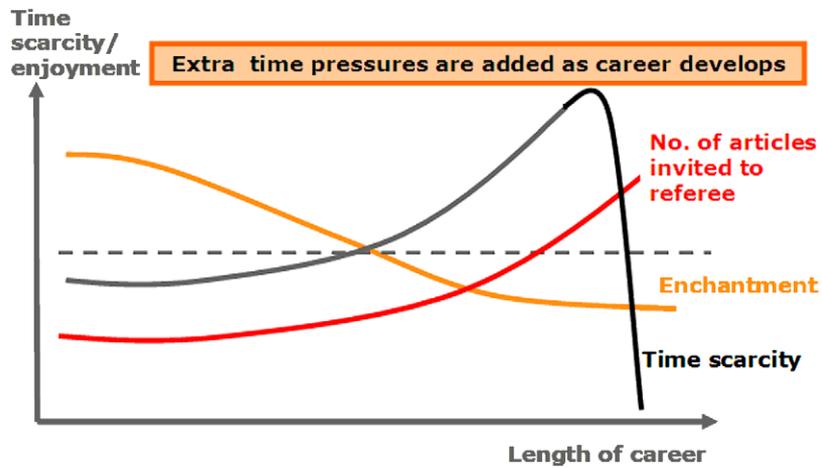


Fig. 1.

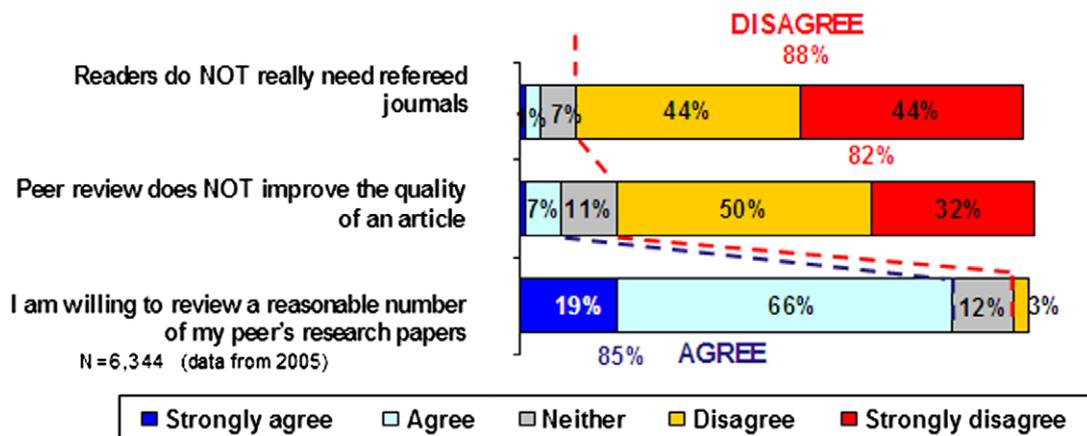


Fig. 2.

provided the number of papers is “reasonable”. Moreover, they believe that peer review actually improves the quality of a paper (see Fig. 2). Since the beginning of 2007 we have conducted follow up work to this study and respondents’ attitudes remain unchanged.

Criteria for accepting a review

Assuring the quality of research in a journal is dependent upon attracting and engaging the best reviewers, so what are the influences acting upon a reviewer when deciding whether or not to accept an invitation to review? There are several critical factors: paper relevancy, journal reputation, the quality of the article and impact factor of the journal. The flip side of the quality influences is that less known journals often struggle to find reviewers. A less obvious factor, but just as critical especially for less well known journals is editorial influence. Editors are very good at exploiting relationships with reviewers and encouraging them to review manuscripts. Reviewers often times are bound to an individual editor because they were former colleagues or friends.

The attributes below give insight into the criteria reviewers use when deciding whether or not to review an article.

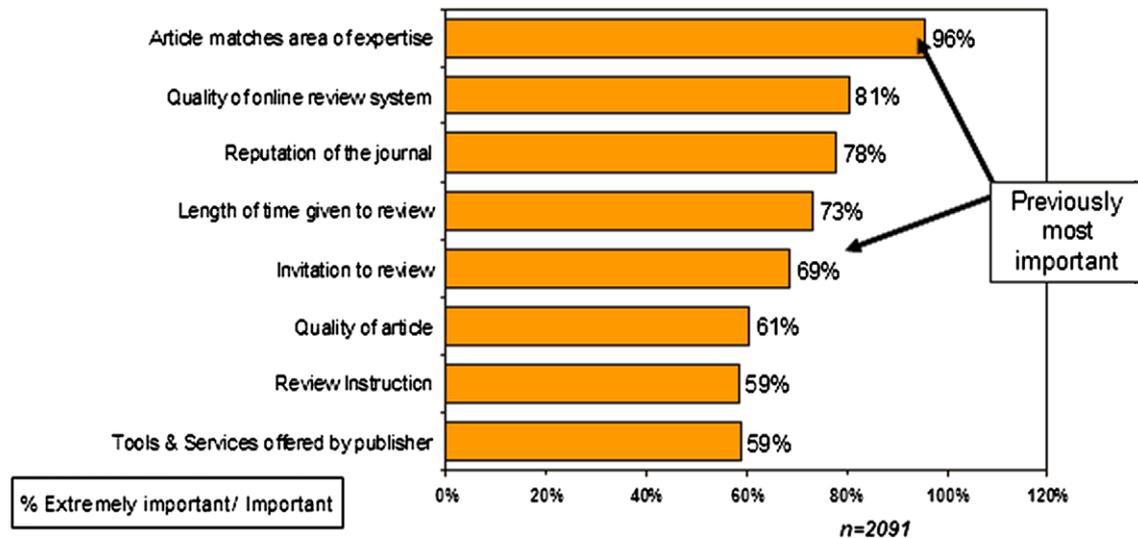


Fig. 3.

Over the past three or four years, we have noticed that there has been a slight shift in the importance rankings of criteria for accepting a review invitation. Relevancy still ranks as most important, but the quality of the online review system now ranks second. The quality of the paper has dropped to sixth position indicating that easy reviewing conditions is taking precedence over quality (see Fig. 3).

Barriers to review

In spite of the widespread acceptance of peer review there are barriers to the process of review. Lack of relevance, being overburdened with large numbers of reviews, and lack of time (typically 2–3 days of work), are the key factors identified for not reviewing. The poor quality of manuscripts, both in terms of the English (especially from Chinese authors) and the scientific quality are also barriers (see Fig. 4). Tackling this issue has an impact upon how we define peer review. Should it start when the editor receives the manuscript, or when it is sent to the reviewer? Reviewers would like to receive reasonable quality manuscripts with a minimum standard of English. If the editor evaluates and rejects manuscript on the basis of English or minimum standard of research is this still peer review. Moreover, if a third party performs technical screening on behalf of the editor should this be considered a formal part of the review process? There is no simple answer.

Other marginal factors included lack of guidance for the reviewer (especially for new reviewers), and insufficient time to decide to accept. Though not measured in 2007, qualitative research from 2004 showed that those driven by their career were less inclined to repeat review. Having completed a previous review for a prestigious journal, as they already have that prestigious name on their CV, there is little reason for them to complete another review.

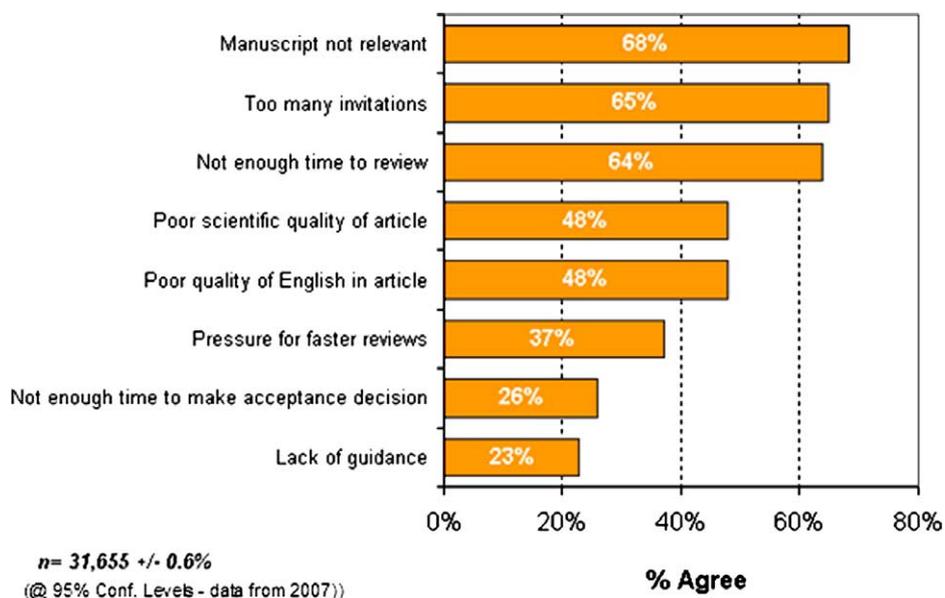


Fig. 4.

Issues in peer review

A number of concerns have emerged in recent year that have been eroding the value of peer review. Peer review is no longer regarded as sacrosanct. It is perceived as increasing the burden on an ever diminishing number of reviewers. Conflicts of interests are considered by some as becoming more prevalent; in an increasingly competitive world it is believed that competitors for funds are writing damning review reports. Others question the choice of reviewer and raise the issue of bias when the author is asked to submit a list of potential reviewers. There is a feeling that the peer review process is too slow, and not rigorous enough, that peer review maintains the status quo, and the process falters when confronted with new theories or alternative methods. Bias is also a concern especially when prestigious institutes are favoured over lesser known institutes, or papers from developed countries are preferred to those countries from Africa or Indian sub-continent. Finally, fraud is considered an issue, by a number within research, and the wider general public. There have been a number of high profile examples of fraud in the last several years notably in September 2002, when Jan Hendrik Schon, tipped to be a Nobel Prize winner, was discovered to have published a series of fraudulent papers [2].

Quantifying the attitudes of researchers in regard to such perceived shortcomings in peer review is problematic due to the nebulous nature of the issues. However, in early 2007 we measured opinion in regard to several key concerns. The research shows that concerns are not as widespread as the hyperbole suggests. Few believe that untruthful reviews are written, though slightly more believe that reviewers unnecessarily delay reviewing to gain competitor advantage. Only a minority believe there is too much plagiarism or fraud. Even here we need to be mindful of context, one reviewer put it quite aptly, “Any fraud is too much – I just don’t believe it’s terribly widespread”.

In regard to the issue of colleagues or friends being reviewers, editors indicate that often only one author nomination will be used, recommendations are more often used on other submissions. Additionally, anecdotal evidence from editors suggests that “friends” of the author are often responsible for writing the most critical reports.

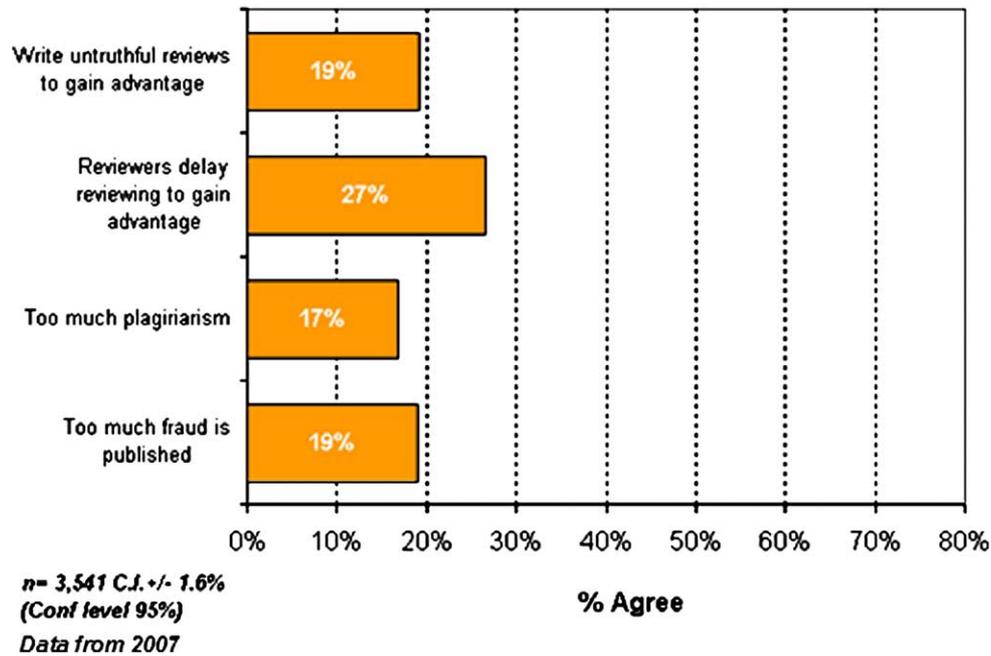


Fig. 5.

Emerging practices in peer review

In recent years new forms of peer review have emerged to tackle some of the perceived weaknesses in peer review. The traditional, and by far the most common, form of peer review is Single Blind Peer Review. This is where the reviewer remains anonymous and the author's identity is known, this applies to the majority of STM journals at Elsevier. The perceived advantage of this approach is that it allows reviewers to write an honest report without fear of recrimination, specifically a negative impact on their careers. Open Peer Review is where the reviewers' names and authors names are known to one another, and often also to the public at large. The perceived advantage is that it prevents plagiarism, encourages honest reviews and discourages malicious ones. This form of peer review emerged within the last the decade and one of the first major journals to adopt it was the *British Medical Journal*. Possibly the most rigorous of peer review, but perhaps least common is Double Blind Peer Review, this is where the identity of the reviewer and author is masked from one another. This form of review is associated with a perception of fairness, since it is easier to give negative feedback and also avoids bias.

Attractiveness of different types of peer review

When it comes to attractiveness of the various type of peer review, reviewers are most in favour of Single Blind and Double Blind – Open Peer Review is a distant third. Double Blind Peer Review is narrowly preferred by most reviewers.

Disadvantages of different forms of peer review

In spite of its popularity Double Blind Peer Review appears to have insurmountable drawbacks. Editors do not embrace it as it requires considerable effort and resource to ensure anonymity, and a num-

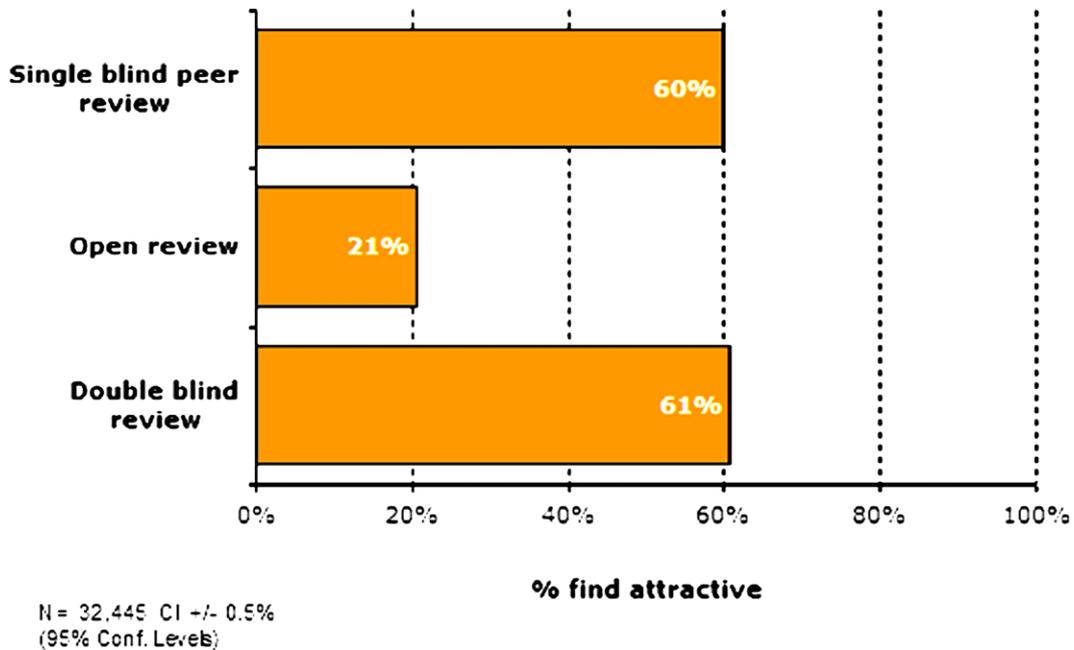


Fig. 6.

ber are unconvinced of its effectiveness. Likewise a number of reviewers remain unconvinced. Mostly because they believe it is impossible to truly obscure the identity of the author. The problem is authors like to cite their own work. When the authors name is not removed from the reference list, an unusual high number of citations or citations to unrelated works will likely reveal the identity of the author. In the opposite situation, when self-citations are removed the lack of the author's name from the reference list can reveal their identity, especially if they are known contributor to the field. Reviewers have indicated that in niche areas, writing style or experiment type can give the author's identity away.

Open Peer Review, was least popular because it was felt that reviewers are less likely to be honest when their names are known, there is an attitude that "closed" peer review is needed to remove poor quality research. For some it was perceived as being too new to generate acceptance, in spite of being around ten years.

It is interesting to note the results of an experiment conducted by *Nature* in 2006. There was a less than enthusiastic reaction from the research community; only 5% of authors agreed to participate and just over half of the articles included received review comments.

New forms of open peer review are being proposed, such as Continuous Peer Review where articles are published on the web and the community at large is given the opportunity to comment on an article. The author is then invited to respond and subsequently revise the manuscript if appropriate. It is just one form of open peer review and one we decided to explore in our research. Initially, responses were very positive, but as researchers became more familiar with the concept they became more questioning, there were concerns about versioning, specifically about being able to reference the 'correct' version of the article. There were also concerns regarding the authority of the persons completing the review, their reliability and the consistency of the remarks.

Concluding thoughts

Many of the issues that surround peer review are less to do with the fundamentals of peer review and more to do with technical improvements to systems or changes to practices at the margins. Overwhelmingly the community value peer review. Attracting and retaining the best reviewers is key to ensuring research articles are evaluated most appropriately. Reviewing is not done for recognition, prestige or gain.

While rewarding reviewers may entice more researchers to review, the reality is that time, relevancy and quality are the key issues to attracting and retaining reviewers. Asking reviewers to write a report only for articles relevant to their expertise is an obvious answer, but problematic to deliver. Developments in submission systems particularly key word databases will help tackle this issue. Reducing the burden on reviewers is paramount, dealing with the issue of relevancy will help resolve this, but inviting reviewers to review first, rather than sending a full paper and simply expecting a report which may or may not be used (a practice on some journals) is the ideal. Technically screening manuscripts will help reduce the burden on reviewers and editors alike.

As we have seen a frequent complaint is that peer review is too slow – migrating to online systems will speed up the process. This has been the experience on a number of Elsevier journals.

Tackling issues of bias might be accommodated by moving to a different form of review, but a more step-wise approach might be advisable. Sharing of results between reviewers could alleviate bias, delays and damning reports. Increasing confidence and understanding in peer review will mitigate some concerns, and this is best achieved by establishing a transparent framework. Transparency benefits readers, authors, reviewers and editors alike, readers know the rigour of peer review, authors can make an informed decision as to where to submit and reviewers have a clearer idea of what is expected of them.

While it is a misperception that fraud is common and unrealistic to expect peer review to prevent fraudulent research being published, it is possible the provision of the author's raw data and sign-off by authors as well as the relevant Heads of laboratories attesting the authenticity of a paper might deter the fraudster.

Good peer review is critical to the exchange of scholarly information, and while the fundamentals are constant, it is nonetheless important to experiment and try different enhancements and various models of peer review. Elsevier is currently testing different forms of peer review on some journals, including open peer review and double blind peer review. It is too early to tell how successful such changes are likely to be. However, it is clear that only by experimenting that we can innovate, and help improve a system so critical to the progress of human endeavour.

References

- [1] J. Barnes, Proof and the syllogism, in: *Aristotle on Science: The "Posterior Analytics"*, E. Berti, ed., 1981, pp. 17–59.
- [2] E. Lerner, Fraud shows peer-review flaw, *The Industrialist Physicist* **8**(6) (2003), 12–17.
- [3] C.M. Olson, Peer review of the biomedical literature, *American Journal of Emergency Medicine* **8**(4) (1990), 356–358.
- [4] J.M. Ziman, *Public Knowledge: An Essay Concerning the Social Development of Science*, Cambridge University Press, London, 1968.
- [5] H. Zuckerman and R.K. Merton, Patterns of evaluation in science: institutionalisation, structure and functions of the referee system, *Minerva* **9**(1) (1971), 66–100.

Web tools for peer reviewers . . . and everyone

Richard Akerman

CISTI

This presentation will explore the peer review and the tremendous transformations brought about by the digital and network revolution, the new kinds of metrics to be used and investigated, the challenges of certification and rewards. Finally, some web tools will be presented.

Today, despite all the technological changes, the main motivation of researchers remains recognition by their peers and showing up in the literature. CISTI, as a publisher and a library, believes that peer review, the lynchpin of the scientific system, has a tremendous core enduring value and has remained basically unchanged. Simple automation of the review process, such as email notices, citation linking and plagiarism detection through automated textual comparisons, can greatly ease the peer review process. Also more sophisticated automation could help locate reviewers and ensure that relevant papers are sent to relevant reviewers, and provide guidance for new reviewers. The question today is whether a peer review system based on the scientists' goodwill is sustainable when scientists are overwhelmed by the journal and submission explosion. Furthermore, their review work lacks recognition and this could be improved by publishing global reviewers' lists and the comments after an embargo period.

During the past few decades, the scientific communication system has been totally transformed and nowadays, a whole library can fit on a chip in a cell phone that can be connected to the Internet. Overall, the scientists' motivations and the peer review process have not changed, but the way information is communicated and used has undergone tremendous transformations, and this is not without some problems. First, there is a problem space to be explored, and a group at CISTI has been trying to understand this new world of digital information and the areas that can be usefully addressed. There are two angles and two directions to the problem space. One might want to either explore the journal space to find existing articles, experts or article clusters (article discovery) or discover ideas, concepts and relationships not clearly visible in the current literature (knowledge discovery). This can be done either retrospectively to find the existing body of scientific knowledge, the citation web and the known relationships between authors (co-citations), or prospectively to discover in real-time new ideas and concepts, not to mention scientists that are not yet connected to the citation web and could be important for science and the nation. Retrospective finding usually starts with a nucleus (an idea, a collection of articles) to find a network of related objects. This can be done through the "recommender services" that use metrics and features to move the initial results into a feature space by establishing closeness and similarity.

As far as features and quality metrics are concerned, some of the quality metrics are networks of citations and authors, groups and projects, and certification (peer review) and the positions in the feature network (above all citations, authors, and journals). Looking retrospectively implies looking at a frozen body of information in the published literature (the dead trees in the forest). But, before exploring the whole network, it is important to make explicit the metaphors used in scientific communication since they might limit and shape the way we think about scientific communication. The term "paper" is misleading since most of the time it just a PDF that is displayed on a screen and that might never be printed. Information "mining" is also interesting. Mining, like in gold mining, means sifting through tones of unrefined material to find a tiny amount of gold. Whereas, literature mining means sifting through highly-refined material to extract some information. Once we explicit the mining metaphor, we see that running mining algorithms on refined information has its value but is often done in the wrong layer (abstracts and full-text) while it should go beyond that layer down to the datasets and full semantic knowledge underlying the paper, and even down to the scientists themselves. Mining also brings up

the question of mining rights, and those are quite limited since publishers are very reluctant to provide access to the fulltext, and this is one of the drivers of the Open Access movement. Improving the mining rights issue will go a long way to clarify the discussions around Open Access and the value of mining since, most of the time, “legal” access to article sets is quite limited. Also, trying to extract concepts by scanning the literature leads to very difficult areas of machine reasoning.

Nowadays, it is possible to go beyond mining and explore the World Wide Web, including repositories and preprints, blogs, wikis, videos, and datasets. Retrospective metrics that could be used for this are either weak or non-existent. So the question is what metrics to use to explore this new space and find valid, real quality science in this chaos of the web. Among the new metrics that can be used are derived reputation (e.g. *Nature* blogs); page rank (as in Google) or valuing pages based on their citations; bookmarks to see how often people have saved links; intentional ranking systems such as explicit thumbs up or down or 1–5 rankings; comments, which presently are hard to obtain but this might change with Web 2.0 savvy young scientists; usage space (page hits, viewing time) and derived quality i.e. trusting the authority of a person or a group. However, all these tracking methods have important privacy and knowledge divulgation implications.

Other challenges are versioning and certification. Metrics should be measuring the usage (downloads and views) of all copies of an article (pre- and post-publication). Some interesting work has been done on versioning in the UK¹ and *eprintweb* is linking preprint to published versions whenever possible. Another thing that might skew metrics is search engines spidering a site and showing robot rather than human interest traffic. The challenge with certification is what should be certified. The imprimatur of peer review should be extended beyond the journal level to cover not only peer-reviewed articles in repositories that are outside of the journal container, but also to all kinds of new web objects such as blogs, wikis, datasets and videos. Scientific work today can be in three-dimensional visualizations that need to be squashed into a two-dimensional text to be explained, whereas in the digital environment the full richness of the datasets can be explored. In terms of digital certification, right now digital signature seems to be the only solution. Certification also relates to award. A new scientist working primarily through a blog is not likely to make it to senior scientist. Rewards need to be attached to these new communication methods in the same way that impact factors are attached to journals. If some of these methods have value, they should have an associated reward.

As a conclusion five websites are presented. The last three might be lighter weight but they also point out opportunities and challenges.

- *Eigenfactor.org*.² The eigenfactor is a different approach to measuring impact other than the trademarked ISI impact factor. It runs algorithms over networks, relies heavily on mathematics, is very complex and not too transparent, except maybe to a mathematician.
- *eprintweb.org*.³ They are working with arXiv data, the arXiv.org repository being most popular with physicists and mathematicians. *eprintweb* runs algorithms to link, wherever possible, the preprints to the final version which is a tremendous way towards solving the version issue.
- *ScienceBlogs*⁴ is an aggregator for scientific blogging but the metric used is often of entertainment rather than scientific value.

¹For the VERSIONS project, see: <http://www.lse.ac.uk/library/versions>.

²<http://www.eigenfactor.org>.

³<http://www.eprintweb.org>.

⁴<http://scienceblogs.com>.

- *Postgenomic*⁵ is another tool for exploring bookmarks and web links to reputable peer reviewed literature, covering all subjects in science, but this a space that is not yet rich enough to explore.
- *ChemRank*⁶ is a site where you can thumb up and thumb down articles and hope that the most valuable scientific content will come up to the surface.

Impact of research assessment on scientific publication in Earth sciences

Bruno Granier

Université de Brest

The impact factor has been for some time a simple way to rank scientific journals. However, it is often misused by being an unyielding yardstick for hiring, tenure and grants. Some reviewers tend to give higher ratings to papers submitted to high-impact journals and articles might not be considered because they are published in journals with a low impact factor or no impact factor at all. There are many players involved in the field, but whether you are a researcher, an institution or a publisher, the main objective is visibility and this visibility can be attained in various ways. This presentation is based on the speaker's broad personal experience and his points of view. After a PhD in Stratigraphy, he worked both on the academic and the industrial sides of the scientific field. In 2002, he started *Carnets de Géologie – Notebooks on Geology*,⁷ an Open Access (OA) electronic journal focusing on Geology, Paleontology and Stratigraphy. He is currently the publisher and chief editor of this journal and for three years now a full professor in Paleontology and Sedimentary Geology at the Université de Brest. The points of view of the different actors are presented in what turns out to be a somewhat schizophrenic scheme.

First, from the author standpoint, authors submitting a paper do not have the same objectives depending on whether they are looking for a position in academe, seeking a job in industry, working as a consultant or looking for career development opportunities. There is no doubt that for all authors, the delay between article submission and publication is an important element. However, for authors looking for work in industry or as consultants, short publication delays are very important, and the impact factor is of little or no importance since the published articles advertise the expertise and know-how of the company that employs them. On the other hand, authors looking for positions or promotions in academe are less concerned with publishing delays since they want to be published in high ranking journals where editorial processes are likely to be longer. However, regardless of the objectives, there are authors and authors. Since visibility and/or impact are the crux of the matter, and to increase said visibility, some authors might resort to intentional or unintentional fraudulent practices. These are known, and the list is not exhaustive, as “slicing”, or *saucissonnage* in French, where one long article is cut into three shorter articles; “bogus signatures” where associate authors are added in hope of reciprocity; “old boy tip” where author names appear in alternate positions on several related but not quite identical papers; “selective or inaccurate quoting” where the only parts of a text that are quoted are those that do not contradict the author's theory; “self-citation or auto-plagiarism” where authors cite any of their articles whether or not they are relevant; “cutting and pasting” where large excerpts of previous publications are

⁵<http://www.postgenomic.com>.

⁶ChemRank has been replaced since with this site: <http://www.hotcites.com>.

⁷<http://paleopolis.rediris.es/cg/uk-index.html>.

inserted in an article without any reference to their previous existence; and “lifting” where sequences of words, data or images from another author publication are deliberately copied and not identified.

Publication duplication can take many forms. Duplication might be intentional such as the same paper given at two different conferences in two different languages and then published in two non-refereed journals, simply to further one’s career in industry. It might also be unintentional when for various reasons a peer reviewed version and a non evaluated version of an article appear in two different journals. Finally, a paper of seminal value might be published some years later in a revised version in order to reach a larger audience. Bogus signatures can be revealed by an increase in the number of coauthors over the years, unjustified by the multidisciplinary nature of an article, as was the case of *Bulletin de la Société Géologique de France* for the years 1957–1999. So the way to evaluate people or laboratories would be to introduce some weighted averages for multi–author papers where the first author will count more than the tenth author. As for lifting, a good reviewer can pinpoint figures or images lifted from previous publications to be reused in a new publication. However, it might take a reviewer expert in the field in question to detect fraud or gross errors, and publications that get through might indicate a lack of any peer review process.

From the reviewer standpoint, the reviewer might still be the *sine qua non* of the evaluation process, but reviewing is a rather difficult and unrewarding task and might not be worth the trouble. Reviewing is done on a volunteer basis, is time consuming, usually anonymous to avoid eventual retaliation, and of no value in the reviewer’s personal evaluation process. Collaborative filtering and open peer review are risky for the same reasons. Also reviewers do not have the opportunity to publish a rebuttal to the papers they review.

From an editor standpoint, it is very hard to find an expert reviewer when the paper is highly specialized. The potential reviewer might also be a competitor, which might lead to blocking or delaying the paper, or an advocate or friend of the author, which might lead to laxism. Some reviewers might be tempted to add their own papers to the work under review in order to increase their own impact factor and citation index. Reviewers are most annoyed by practices such as padding references with references to articles that are not relevant and by incorrect references, all of which can highlight the fact that the cited publications were not read. So, as a paleontological reviewer, the speaker observes the following far from perfect rules: any publications without illustrations of significant fossil taxa or relocation in a specific collection will be rejected; references to abstracts will not be accepted; previous papers by the same author will be systematically screened to avoid duplication; articles published by institutions or authors in violation of scientific ethics will be blacklisted according to some authority list; and computer programs designed to detect substantial excerpts of previous work in a new article submitted for review will be used.

From the standpoint of the publisher of *Carnet de Géologie*, an online OA journal, the Google effect is very appealing and 3 million cites can be searched. Google is very good for OA journals since they are easily accessible and consequently citation rate should increase. Some papers are simply not cited because they are difficult to access. It is quite understandable that commercial publishers use new tools such as usage counters to filter which kind of papers are the best in terms of visibility, but not in terms of quality. The impact factor is important for journals, but books, monographs, maps and many electronic OA journals are left aside since some electronic OA publications are too new to have accrued an impact factor. Therefore, a large part of scientific information is not taken into account in the researchers’ evaluations. There is some sort of snowball effect as authors are attracted to high-impact factor journals, so the reservoir of potential authors for journals with low or no impact factors is very low. Another effect of the impact factor is that articles are rather short and targeting high-impact journals such as *Nature*

or *Science*. One might say that publishing is facing a biological-like crisis. Commercial journals are the survivors while entire populations of regional, national, learned society, museum, and university journals are being eradicated because of the use or misuse of metrics such as citations counts. The good thing is that a new category of opportunist journals such as those of BioMed Central and PLoS is emerging.

To conclude, Eugene Garfield, the father of the impact factor, now compares his brainchild to nuclear energy that is to say a force that can help society but can be misused if left uncontrolled. Bibliometrics or no bibliometrics, the goal that remains is to increase visibility, and today visibility is on the Net and on Google. Should the perpetrators of some of the manipulations that were mentioned be equally blamed? Most are just trying to keep afloat and they cannot be blamed when the only way to survive is for journals to increase their impact factor and for authors to increase their citation index. The forthcoming crisis will be the Google effect and it is already at our door.

Faculty of 1000 Medicine: Post-publication peer selection

Pritpal S. Tamber

Faculty of 1000 Medicine

Before presenting Faculty of 1000, a few fundamental points need to be made.

First, there is the issue of data and the difference between data used for a study and data presented when this same study is published. In medical research, and especially in large randomized controlled trials, a huge amount of data is measured but no journal editor is willing to publish all the data for reasons such as space constraints, non relevance to audience, costs or other reasons which have little to do with the communication of science. So, for example, in medical research, up to 120 endpoints can be measured, but to publish a study, only about 17 of them can be reported. This is an incredibly biased way to look at the data because, when only 17 out of 120 endpoints can be published, you lose the context provided by the 103 that were left out. This is why, from a personal standpoint, journal articles need to change rapidly, and why bibliometrics based on published papers lack a certain amount of credibility.

The second fundamental point is peer review. In Medicine, there is a group called International Committee of Medical Journal Editors (ICMJE) representing about 20 large and influential journals, and even their definition of peer review is a little vague. But, even if one cannot be certain about what exactly peer review is, it is important to see what it does. Given that citations in PubMed have shown a 20% increase in one year, which is quite an explosion, but one might wonder why bother with peer review when, according to figures in previous presentations, one fifth of authors feels that peer review does not improve article quality. Furthermore, if 23% of the respondents in the Elsevier study feel that there is a lack of guidance in peer review, one might wonder whether almost a quarter of reviewers have a clue to what they are doing. So, if no one knows what peer review is and how to do it, and one fifth of the people feel that it doesn't add anything to their articles, then peer review is in no way a mechanism to certify quality.

There is quite a fundamental problem if, as a community, we know that printed articles are a poor representation of a full research project and that peer review is a poor mechanism for certification of quality. The classic manifestation of all this is Robert Maxwell, the media tycoon, who made a significant amount of money with publishing scientific journals by turning peer review into a marketing tool. He claimed he'd get peer review away from those closed networks of researchers with biased ideas and unfair views. One might say that this twisting of peer review is akin to what has happened with "evidence-based medicine", as a concept. Both started off as good ideas in the scientific community and then became a strong marketing message.

Anyway, peer review must be considered very carefully. It is far from perfect but according to Liz Wager, a biomedical publications expert, it is the best thing we have. It is not a question of altogether abandoning peer review. However, before pinning all our hopes on it and looking into what we could add to it, we should ask some basic questions on the nature of peer review, on what it does, on how to communicate what it is to the people who do it, on how to enhance the process, and on how to study the alleged improvements it brings to articles since. Indeed, despite three congresses on peer review in the last ten years, there has been little evidence to back its use.

Regardless of the above, there have been interesting innovations in peer review such as the three following examples. First, the example of *Atmospheric Chemistry and Physics (ACP)*⁸ which is ranked as twelfth out of 169 in its field. Submissions are put online for 8 weeks on *ACPDiscussion (ACPD)* and people are invited to comment. Then, the article is removed from *ACPD* and submitted to a formal closed peer review process. If accepted, it is published along with the original discussion comments but, for some odd reason, without the formal peer review reports. At BioMed Central, everything was open and transparent within its medical journals. Reviewers knew who the authors were and vice versa, and when an article was accepted, all the peer reviewers' reports went online next to the article. In 2005, *ACP* published 240 papers but only one in four received comments during *ACPD* and these comments were half a page long on the average, which is rather light when it comes to reviews. The second example is *Biology Direct* (2006),⁹ a BioMed Central journal. It has author-guided peer review which means that, during online submission, authors can select whom they would like as a reviewer and that reviewer then decides whether or not to do the review. For such an open peer review system to work, *Biology Direct* has a very large editorial board, up to 200 persons compared to 30 for a standard editorial board. In this system, the author does a great part of the administrative work but, in the end, the article is published with all the reviewer and author comments and discussions. This experiment deserves to be better publicized since it is quite courageous an experiment. The third example is that of *PLoS ONE*.¹⁰ It seems that this publication is not doing well so far. It publishes articles with a very low barrier to publication and has a very nice online commenting function.

Faculty of 1000 tries to find solutions to the following problems: First, there is the problem of information explosion (too many articles and journals), researchers feel overwhelmed and can no longer keep up with the literature or know what is worth reading. Despite searches running in PubMed and other aggregating services, they worry that they might miss something important, and the problem is far from improving. Amazingly, new journals are coming out all the time. There is a maxim making the rounds in Medicine that says that in biomedical publications “only 5% are methodologically sound, of which only 5% is clinically useful”. The question is how to find that 5% clinically useful material. Citations are of little value in Clinical Medicine since few clinicians write articles, most of them are practicing, and unlike PhDs and postdocs, MDs are barely taught how to read scientific articles. Also the 5% methodological sound research looks at communities that are so strictly defined in terms of population, risks, etc. that it is impossible to apply to patients in general. Often quality and usefulness are not synonymous in clinical medicine where lots of clinical nuggets are discovered through experience rather than research, and most physicians are not very likely to base patient treatments on citations.

When it comes to “peer networks” such as *MySpace* and *Flickr*, there is a general feeling that creating such online communities for physicians enables them to share their knowledge. One of the most serious

⁸<http://www.atmospheric-chemistry-and-physics.net>.

⁹<http://www.biology-direct.com>.

¹⁰<http://www.plosone.org>.

is “*sermo*” in the US, a closed site where you need a US doctor registration number to even log in. However, clinicians being nomadic by nature have little time for online comments and votes. Also, in these types of networks, it is difficult to identify and control conflicts of interest and lobbying by patient advocacy groups, regardless of how justified it is. So all in all, it is preferable that peer networks be moderated somehow. Another closed network is *Faculty of 1000 Medicine*¹¹ with 2,400 experts selecting articles that matter, commenting on them, classifying them according to 1–7 types and rating them on a 3-point scale. It covers 18 medical specialties broken down into 202 sub-specialties. *Faculty of 1000 Medicine* was launched in 2006 while its predecessor, *Faculty of 1000 Biology* has been alive since 2002. One of the beauty of that service is that more than one person can comment on an article, so different specialists can give their point of views on a drug, a treatment, etc. This means that the multidisciplinary aspect of reading papers is reintroduced since a single article can be classified into a number of areas.

Due to lack of time there was no discussion but some issues were raised during the presentation itself focusing briefly on statistical interpretation and more extensively on evidence-based vs. opinion-based medicine.

Interactive Open Access publishing and collaborative peer review for improved scientific communication and quality assurance

Ulrich Pöschl
Max Planck Society

Statistics have been used and abused and this presentation will explore how statistics can back existing viable solutions to improve scientific communication and quality assurance, using the journal *Atmospheric Chemistry and Physics*¹² as an example.

Recent high profile cases of scientific fraud have fuelled the discussion of scientific quality control. A problem of similar, if not greater, importance is the large proportion of carelessly prepared scientific papers that dilute rather than enhance scientific knowledge. Both problems are indicative of shortcomings in the traditional peer review system, yet many scientists and publishers believe that peer review remains the best available approach for scientific quality assurance. Nevertheless, requests for improvements are commonplace. Among the suggestions are the public exchange of referee comments and author replies [1,2] and public discussion of manuscripts before formal publication [3,4]. These concepts and their advantages can be efficiently combined with the strengths of traditional peer review.

Quality assurance of scientific publications usually proceeds through two pathways: a pre-publication short term assessment by designated referees during the peer review process, and a post-publication long term assessment by the scientific community through comments, citations, review articles and monographs. Both can be combined in a collaborative peer review process where members of the scientific community can participate in the assessment of scientific manuscripts through interactive comments, in addition to designated referees’ reports.

Interactive Open Access journal concept

Open Access has many advantages that hopefully publishers will soon recognize, so that a common business model can be found. Free online availability of scientific research publications has scientific,

¹¹<http://www.f1000medicine.com>.

¹²<http://www.atmospheric-chemistry-and-physics.net>.

educational and economic advantages. It enhances research impact and productivity, and improves quality assurance; it opens access to students and the general public, thereby leading to equal opportunities in the information society; and it frees a somewhat convoluted scientific information market from issues concerning subscription and usage, cost and benefit and library budget, enhances efficiency and facilitated innovation by opening new ways to format, distribute, evaluate and archive information.

Open Access (OA) is not a threat to scientific quality assurance and can provide some urgently needed improvements. It is fully compatible with traditional peer review as open access journals such as *PLoS Biology* and *BMC Structural Biology* among others have shown. It gives reviewers wider access to information and fully enables new processes such as collaborative peer review. But OA is not problem-free.

The collaborative peer review approach is pursued by the open access journal *Atmospheric Chemistry and Physics (ACP)* and a growing number of sister journals published by the European Geosciences Union (EGU).¹³ These journals have a two-stage publication process with public peer review and interactive discussion [5,6].

Traditional peer review is insufficient in today's rapidly evolving scientific environment. Editors are overwhelmed by the number of submitted papers and there is a shortage of reviewers able to review highly specialized publications. Closed peer review works well but it entails publication delays, information loss by not publishing interesting comments and allowing for little traditional discussions. For example, in Atmospheric Sciences, the 1978–1998 comment to article ratio has gone from 1 in 20 to 1 in 100.

Scientific publication is faced with a speed versus quality dilemma. Today, information must be disseminated quickly, so rapid publication is needed. The problem is how to reconcile rapidity and the thorough reviewing process needed to identify scientific fraud and duplications. The solution is to use the above-mentioned two-stage open access publication process with collaborative peer review. First, a submitted paper is rapidly published as a discussion paper following pre-selection by editors and/or referees. It is more than a preprint since it is fully citable and permanently archived. At that stage, it is open to the scientific community for interactive public discussion and the comments are published alongside the discussion paper (*Atmospheric Chemistry and Physics Discussions–ACPD*). The second stage is similar to traditional publishing with the final publication of the revised paper. The only real difference is that now the black box of the peer review process is transparent even though the option of anonymity remains. Despite what some might think, interactive OA publishing increases the number of meaningful comments (4–6 average for *ACP*). All in all, interactive OA publishing is a win-win situation for authors, referees and readers. The discussion paper is an opportunity for free speech. Collaborative peer review provides direct feedback and public recognition of high-quality papers. It prevents obstruction and plagiarism, documents scientific flaws, critical comments and controversial arguments, discourages careless or useless papers, and provides for technical corrections, therefore lightening the traditional reviewing and editing workload. All in all, this process guarantees maximum quality of the final paper.

Atmospheric Chemistry and Physics

ACP was established in 2001 and now publishes about 400 papers per year. *ACP* statistics show a monthly average of 40 submissions and a less than one month submission to publication time. On average, one in four papers receives a comment from the scientific community in addition to the comments

¹³<http://www.egu.eu/publications.html>.

from designated referees (in traditional journals this happens for about one in 100 papers). And there are typically 0.5 pages of comments and replies per page of original discussion paper. These cover the full spectrum of opinions – from harsh criticism to open applause – and provide a wealth of additional information and evaluation that is available to everyone. Last year, interactive discussions for these papers in the discussion forum generated 7,000 pages of comments. Today, *ACP* is financially viable thanks to its average 1,000 Euro publication charges.

ACP statistics confirm that collaborative peer review facilitates and enhances quality assurance. The journal has a relatively low overall rejection rate of less than 20%, but only a few years after its launch *ACP* citation statistics for 2006 show a high journal impact factor (4.36) for papers published in 2004 and 2005, and *ACP* was ranked first out of 47 journals in “Meteorology and Atmospheric Sciences”, and one of the highest in “Geosciences, Multidisciplinary” (129 journals) and in “Environmental Sciences” (140 journals).¹⁴ Even though *ACP* tries to tone down the impact factor on its site, it is nevertheless worth mentioning, and according to such metrics, one can say that *ACP* is quite successful. These numbers support the anticipation that public peer review and interactive discussion deter authors from submitting low quality manuscripts and, thus, relieve editors and referees from spending too much time on deficient submissions.

Conclusions and outlook

The European Geosciences Union (EGU) journals are the proof that open access publishing process works. The process was first tested with *ACP*. It got off to a quick start, P.J. Crutzen, a Nobel laureate, is chairing its advisory board and it has captured a good share of the journal market in the field. Even though they might not have a Nobel laureate on their boards, other EGU journals are doing quite well. In partnership with Copernicus, an open access service provider and Max Planck Society spin-off, EGU currently publishes seven interactive OA journals (*Atmospheric Chemistry & Physics*, *Biogeosciences*, *Climate*, *Cryosphere*, *e-Earth*, *Hydrology*, *Ocean Sciences*) and three traditional peer reviewed journals have been converted to Open Access (*Annales Geophysicae*, *Natural Hazards*, *Nonlinear Processes*). EGU has already a solid experience in OA journal publishing and is planning to move more and more in this direction.

Other open access alternatives to traditional publishing already exist. Some journals such as *Journal of Interactive Media in Education*, *BioMed Central Biology Direct* and *British Medical Journal* provide Open Peer Review with no referee anonymity. Other journals such as *BioMed Central Medical Journals*, *Behavioral & Brain Sciences* provide pre-publication history and peer commentaries, but do not integrate peer review and public discussion.

Future perspectives focus on finding an efficient and flexible combination of new and traditional forms of review and publication through multiple stages and levels of interactive publishing and commenting, statistical analysis and quality assurance feedback, and integration in large-scale open access publishing systems using repositories and “assessment houses” instead of journals that provide discussion forums for public peer review and interactive discussion. Feedback loops of quality assurance between discussion papers and final papers can prevent the rejection of discussion papers that were in fact better than the final papers.

In the future, two assessment styles, community assessment and organized analysis, could co-exist since they address different needs. Community assessment, with commentaries, review articles, citation analysis (Open Access opens huge possibilities here), mean that the assessment process is slower

¹⁴http://www.atmospheric-chemistry-and-physics.net/news_acp_isi_impact_factor.pdf.

but more accurate in the long term, while organized analysis is immediate but cruder. The scholarly communication system needs to be disaggregated, open to current agents, new entrants, value added services and various business models, and to give a place to certification through public peer review and interactive discussion forums. The vision is to use Open Access and Collaborative Review to promote scientific and societal progress in a context of global information commons [7]. Access to high-quality scientific publications will provide scientists and society with greater volumes of quality information. Documentation of scientific discussions will evidence controversial opinions and open them to questions. Finally, increased transparency and rationalism will be a role model for the political decision process. This vision could be achieved by promoting OA publishing, emphasizing quality assurance and interactivity, and improving scientific evaluation and rating methods. Open Access can be promoted by requiring public access to publicly funded research results and by transferring funds from subscription to OA publications. Some traditional publishers already provide OA options. The transition to Open Access is where the problem lies. Library subscription budgets could be cut by 10–30% a year and the funds thus collected used to finance an OA publishing system. Access to repositories and self-archiving is not enough and qualified multi-layered information is needed. Some journals besides EGU's are already going in this direction (*economics e-journal*, *PLoS ONE*). Scientific evaluation and rating methods could be improved by evaluating individual articles rather than individual journals and by refining statistical parameters for citation, download and usage through interactive commenting and rating.

Moreover, public review, discussion, and documentation of the scientific discourse can serve as an example for rational and transparent procedures of settling complex questions, problems, and disputes, i.e. as a model for the further development of the structures, mechanisms, and processes of communication and decision making in society and politics [8,9].

References

- [1] D.K. Lahiri, Discourse among referees and editors would help, *Nature* **439** (2006), 784; doi:10.1038/439784d.
- [2] D. Kennedy, Nick Cozzarelli, *Science* **312** (2006), 159; doi: 10.1126/science.1128103.
- [3] E.G. Van Meir, Opening the chamber of peer-review secrets, *Nature* **429** (2004), 803; doi: 10.1038/429803c.
- [4] L.B. de Carvalho, Pressure also leads to worthless publications, *Nature* **439** (2006), 784; doi:10.1038/439784a.
- [5] U. Pöschl, Interactive journal concept for improved scientific publishing and quality assurance, *Learned Publishing* **17** (2004), 105.
- [6] T. Koop and U. Pöschl, An open, two-stage peer review journal, *Nature Web Debate on Peer Review*, 2006.
- [7] P.A. David and P.F. Uhler, *Creating the information commons for e-Science*, Paris, 2005.
- [8] U. Pöschl, *Open Access & Collaborative Peer Review*, DFG-iFQ Working Paper No.1, 2006, 43.
- [9] E. Bodenschatz and U. Pöschl, Qualitätssicherung bei Open Access, in: *Open Access Handbuch*, B. Malina ed., UNESCO Kommission Deutschland, Bonn, 2007.