Maintaining Live Discussion in Two-Stage Open Peer Review

Erik Sandewall

Linköping University Post Print

N.B.: When citing this work, cite the original article.

Maintaining live discussion in two-stage open peer review

Erik Sandewall[1,2]*

[1] Department of Computer and Information Science, Linköping University, Linköping, Sweden
[2] Department of Publication Infrastructure, Royal Institute of Technology, Stockholm, Sweden

Open peer review has been proposed for a number of reasons, in particular, for increasing the transparency of the article selection process for a journal, and for obtaining a broader basis for feedback to the authors and for the acceptance decision. The review discussion may also in itself have a value for the research community. These goals rely on the existence of a lively review discussion, but several experiments with open-process peer review in recent years have encountered the problem of faltering review discussions. The present article addresses the question of how lively review discussion may be fostered by relating the experience of the journal Electronic Transactions on Artificial Intelligence (ETAI) which was an early experiment with open peer review. Factors influencing the discussion activity are identified. It is observed that it is more difficult to obtain lively discussion when the number of contributed articles increases, which implies difficulties for scaling up the open peer review model. Suggestions are made for how this difficulty may be overcome.

Keywords: open peer review, community peer review, two-stage peer review, live discussion

1. INTRODUCTION

Open peer review has been proposed for a number of reasons, in particular, for increasing the transparency of the article selection process for a journal, and for obtaining a broader basis both for feedback to the authors, and for the acceptance decision. It has also been proposed that the contents of the reviewers' comments and of the authors' responses to them may in themselves be of interest to the community of researchers in the area of the work, and that they should therefore be published and preserved.

Several of these goals rely on the existence of a lively review discussion. If the discussion falters then only the transparency goal remains, and if the discussion is limited to comments by two or three appointed referees and the authors' responses to them then the review process is little more than traditional peer review where merely the reviews are made publicly available.

Unfortunately, several experiments with open-process peer review in recent years have encountered the problem of faltering review discussions, for example, the experiment made by *Nature* in 2006 (Editorial Report: Nature's Peer Review Trial, 2006). It is therefore of interest to study examples of open peer review where it has been possible to maintain lively discussion, at least in some parts of the experiment, and to discuss the factors that may affect the volume and the character of the discussion.

The Electronic Transactions on Artificial Intelligence (ETAI) was an early experiment with the use of an open peer review process where lively review discussion was an explicit goal, and in fact an essential ingredient in the journal's review process. This journal was started by myself in 1997 because of my dissatisfaction with traditional peer review, and with an idea about an alternative peer review method that would not suffer from the same problems. Some parts of the journal's activities enjoyed lively review discussions; other parts did not. In this article I shall describe the experience from the ETAI in this respect and compare them with observations of one other two-stage peer review journal. I shall observe that the problem of maintaining liveliness seems to be related to the question of scaling up of the journal's size, and conclude with suggestions for how scaling up may be achieved without sacrificing liveliness.

2. RATIONALE AND CONSTITUENCY FOR THE ETAI

Around years 1995 and 1996 I was concerned about the following problems with traditional, confidential peer review:

- The process can be manipulated. This is bad in itself, and it inspires distrust.
- If an article is rejected although its contents actually merit publication and this is discovered some years later, it is in practice impossible to correct the mistake and give due credit to the author. This is always damaging, and in particular so for articles that are ahead of their time.
- If an article is controversial, then the controversy should be brought out in the open so that everyone can make his or her own opinion about it. It should not be kept inside the close walls of the peer review process.
- Since reviewers are anonymous, they can not get proper credit for the work they put in. Quality control of the reviews is difficult for the same reason.
- Peer review is intended to serve two purposes: to provide feedback to the authors so as to improve the article, and to give a guarantee of quality. Its efficiency with respect to the first aspect is often marginal and could be improved.

Considerations similar to these have been discussed by many authors both before and after that time; see for example Gura (2002) and Benos et al. (2007). They led me to propose and to start the *Electronic Transactions on Artificial Intelligence* (ETAI)[1]

---

[1]http://www.etaij.org/

as an attempt toward the solution of these problems, without losing the strong points of conventional peer review. The research area being addressed by the ETAI is Artificial Intelligence, and some background about the character of this field is relevant for understanding the development of the ETAI itself.

Artificial Intelligence is a relatively independent branch of computer science that has strong connections to formal logic, formal linguistics, cognitive science, and a variety of other disciplines ranging from control engineering to psychology. The social structure of this partly interdisciplinary field of research is relevant for the ETAI peer review model: artificial intelligence can be viewed as consisting of a fairly large number of specialities, each with its own "college" of researchers that are active in the area, that meet regularly at conferences and workshops, and that to a large extent know about each others' research directions. Each "college" has a worldwide membership that may count one or a few hundred researchers including the graduate students. The likely readers and the likely peer reviewers of a research article are usually found in the circuit of such a "college."

Structures of this kind occur in many scientific disciplines but apparently not in all.

A second, important consideration concerns the character of research in the field. There is a combination of theoretical research and systems-building research. Theoretical research is done with standard methods of applied mathematics as applied to formal logic. Systems-building research is often done in large projects involving many participants over an extended period of time. It is generally acknowledged in the field that the results of systems-building research do not easily conform to the conventional publication formats, since it is difficult to identify "result modules" that are sufficiently independent of the rest of the large project and that can easily be published. Also, even if it is possible to construct a number of such "result modules" from a large project, the collection of these often fails to give a correct insight into the real results of the entire project. Finally, a large part of the real project results have such a character that they can best be communicated in a dialog-like setting where the pros and cons of different design decisions, for example, can be presented and discussed. They therefore do not fit so well into a framework where one expects to publish definite and unchallengeable results.

3. CONCEPTS AND DISTINCTIONS

The concept of "open peer review" is presently being used for several fairly different models of peer review. A basic distinction can be made between *open-names peer review* which is similar to traditional peer review except that the identity of the reviewers is shown openly, and *open-process peer review* where interested parties are invited to join the peer review process that takes place before an article is accepted for a journal or other similar venue. Hodkinson (2007) uses the term *community peer review* for open-process peer review and introduces additional distinctions.

One may notice that open-names and open-process approaches may be combined in several ways, so that one may use open-process peer review that does not operate with open-names, and vice versa. The present article will only address open-process peer review and will use the term *open peer review* as a synonym.

The ETAI used a *two-stage* peer review process (Sandewall, 1997b, 2006, 2009) that is based on both open-names and open-process, and that works in the following steps. Submitted articles are screened for relevance and if they pass this filter, they are posted on the journal's webpage and made available to the community of researchers in the research area that the article addresses. This begins a 3-month period of open, constructive critique: questions are posed to the author, objections can be made and answered, and so forth. This *review* process is entirely open, so the names of all participants are seen openly (with rare exceptions). After the open discussion period, the author is able to revise the manuscript based on the feedback obtained, and resubmit it to the journal. It is then sent for *refereeing* to two or three referees whose identity is not divulged. The task of the referees is to only make a pass/fail decision, and they are asked not to propose additional changes in the article.

This separation of the peer review process into two stages reflects the two major purposes of peer review, namely, to improve the quality of submitted manuscripts, and to establish quality standard. Conventional peer review integrates these two goals, whereas in our system they are separated so that the purpose of the first stage is only for feedback to the author and for quality improvement, and the second stage is only for maintaining the quality standard. Therefore there is only one revision of an article, namely between the first and the second stage, and the version of the article that is submitted to the second stage becomes the final article if it is accepted. (This is the principle, but in fact there were occasional exceptions where a second round of minor revisions were requested).

The concept of *publication* needs to be made precise in the context of open-process peer review, in particular because of the very peculiar way that this word is used in the scientific community. The original and natural meaning of "publication," in the sense of an activity, is of course to "make public." However, in the context of scientific communication it is often considered to mean "published after having been accepted in a peer review process." This terminology is problematic for us since open-process peer review requires by definition that articles are made available to the scientific community in its topic area for the purpose of *starting* the peer review process.

It is interesting to notice how this peculiar terminology has arisen in the first place. It can be led back to the establishment of the *Ingelfinger rule* (see, e.g., Angell and Kassirer, 1991), a principle developed by Franz Ingelfinger in the 1950s for use in the editorial offices of the New England Journal of Medicine, stating that this journal would not publish any articles whose contents were also published elsewhere, and requiring authors of submitted manuscripts to abide by this rule. The effect of this rule was to establish the journal as an archival one: if it is intended that annual volumes of a scientific journal are to be preserved in university libraries then it is inefficient to store several copies of the same article, whereas for journals that are received, read, and discarded this is not much of an issue.

The Ingelfinger rule was quickly adopted by many other journals at the time and has remained popular. Unfortunately however it was established only a few years before the spread of affordable small-scale reproduction technology using mimeograph

machines, and later on using large-volume copying machines. These had the effect that researchers in some fields started to prepare "departmental reports" for distribution to peers ahead of journal publication.

Journal editors and publishers reacted to this technical development in two different ways. In some areas, such as medicine, it was correctly observed that such departmental reports were publications, and according to the established rule the existence of such a report precluded publication of the same results in a journal, which of course effectively prevented the practice from being adopted at all. In other fields, such as mathematics, physics, and computer science, it was decided instead that departmental reports was a valuable thing to have, but instead of retracting the Ingelfinger rule one decided that a departmental report would not count as a *publication,* thereby making it possible for journals to accept such manuscripts. It is this game with the words that haunts us today.

This was an important issue when the ETAI was launched, in particular since one of the critical questions that we heard when we presented our novel peer review model was: *if an article is distributed openly before being accepted to the journal, how can one avoid that someone else "steals" the results and publishes them in his own name?* There was only one way of addressing this problem, namely, to return to the original meaning of "to publish" and to state as a terminological policy that an article was to be considered as published exactly when it was made public to the members of its peer community, which meant, well before it was accepted to the journal, and without any guarantee whatsoever that it would eventually be accepted. In this way, the priority for the results in the article should count from the date when the article was first made available.

This policy immediately led to a second question: if the article was published before being accepted for the journal, then who was the publisher? This led to the creation of the *Linköping University Electronic Press* (LiU E-Press)[2] as an open access publisher precisely for the purpose of having a publisher for submitted articles.

Consequently, whereas the Ingelfinger rule says that the journal will not publish previously published articles, our procedure implied that the journal would *only* publish previously published articles, namely, after the successful peer review of an article that had been published so that it could be peer reviewed.

These considerations concerning the concepts of publication and of publisher were laid out in an article that was published by the LiU E-Press in 1997 (Sandewall, 1997a). It was of course important to obtain as broad acceptance as possible of these unconventional ideas. I was therefore glad to have been invited to a working group that had been asked by the Association of STM Publishers (Science-Technology-Medicine) to find an answer to the question: *What should be considered as a publication in the electronic age?* – the problem being of course that there is no obvious original copy of a document that is produced and disseminated electronically.

The working group's report (Frankel et al., 2000) reflects some of the ideas that have been described here, in particular insofar

as it recognizes several successive versions of a publication, where the peer reviewed version is designated as "final" but the earlier versions are also recognized as "publications."

However, in my opinion the group never answered the basic question that had been posed, that is, how do you define the *publication* then? My own answer to this question was and is that one must first define an *electronic publisher* as an organization that is able to organize, preserve, and disseminate electronic documents persistently, and then define an *electronic publication* as an item that has been published by such an electronic publisher. The group did not however want to address this admittedly somewhat philosophical issue.

4. CHALLENGES FOR A NEW PEER REVIEW MODEL

ETAI's two-stage, open-process peer review model was easily accepted in its own research community of Artificial Intelligence. It was given particular strength since we secured the support of two important parties: it was published under the auspices of the Swedish Academy of Sciences and of the European Coordinating Committee for Artificial Intelligence, which is a federation of national A.I. societies.

This does not mean that everything was easy. The challenges were of several kinds:

- Doubts about the model by representatives of other disciplines, which in turn caused some of our colleagues to stay away from it.
- The problem of getting the flow of submissions to start initially.
- The problem of maintaining coherence in a journal that was divided between a number of specialized areas.
- Insufficiency of the computational and administrative infrastructure.

Any new journal of this kind is likely to face these questions, and it is important to be clear how a particular model for open peer review can handle them. I shall discuss them in turn.

4.1. DOUBTS ABOUT THE OPEN-PROCESS PEER REVIEW MODEL

A number of persons told us that the ETAI peer review model simply would not work when it was first explained to them. Their pessimistic predictions turned out to be incorrect. It is interesting to note that the reason for the incorrect predictions was because people extrapolated from their acquaintance with traditional peer review but the extrapolation was not applicable.

In particular, one objection was that the model would not work since no one was going to contribute critical comments to the open peer review discussion for not risking to make enemies with the authors. This analysis was incorrect because whereas a critical comment in conventional peer review is to the author's disadvantage (at least in an immediate sense), in the two-stage peer review scheme the author has a fair chance to respond to the critique, and also to make a correction in the article if this is warranted.

In fact, several of our authors reported that they *were glad* to receive critical comments since this made the discussion more lively, and therefore they obtained more attention for their article. This is like at a Ph.D. defense: a dull session is not appreciated, and the best is if the candidate obtains difficult questions and is able to answer them well.

---

[2]http://www.ep.liu.se/

Another objection was that we would be overwhelmed by an avalanche of so-called "junk" articles, since authors would see a chance to have their articles published without peer review. This did not happen exactly because of the openness in the system. Under the conventional peer review scheme it does not "cost" anything to submit a substandard article since only the reviewers will know. In our model the quality of the article and the fact that it was not eventually accepted would be clear to everyone.

Predictions of this kind have appeared repeatedly, e.g., in an editorial of Editorial: Revolutionizing Peer Review (2005), but repeated practical experiences seem to refute it. The experience of the journal *Atmospheric Chemistry and Physics* (ACP) is similar to ours in this respect (Koop and Pöschl, 2006).

A complementary prediction was that we would not receive any submissions at all since no one would want to risk the shame of not having their article accepted. Fortunately it turned out that authors were more wise than that. We did decline some contributions and this did not have any noticeable effect on the flow of contributions afterward. Conversations with actual and would be authors suggested that this was not perceived as a problem.

Another objection concerned rejected articles. An article that has been rejected from a journal that uses conventional peer review can be submitted to another journal, but in our case this might not be possible, it was argued, since the article has been published in the formal sense. This did not seem to be a problem in practice, however, in particular since Computer Science is an area where prepublication using departmental reports is widely used and accepted, so journals tend to be generous in their interpretation of "previously published." It might have been different in another field.

However, it should also be said that the practice where an author of a rejected paper resubmits the same paper to another journal without first acting on the reviewer feedback, is in fact a problem for the research publication system. Under the ETAI system it is still possible to submit repeatedly in this way, unless the second journal has a principle against it, but since the negative reviews from the ETAI are publicly available the author will have strong incentives to address the critique before the new submission.

Yet another objection was that the delay of 3 months until the acceptance of an article in the journal was too long. In the AAAS/UNESCO/ICSU workshop in 1998, Parker (1998) of the Royal Society of Chemistry stated[3]:

> [This] contribution describes a very nice refinement to open review. However, I think most chemists would be horrified by the thought of peer review taking three months for the initial phase plus a bit longer for the intensive phase. The current average time from receipt to publication in RSC's flagship journal, Chemical Communications, is under 80 days and decreasing! I think this raises the distinct possibility of divergence of peer review policy among disciplines.

and later on:

> Perhaps chemistry is less contentious and results less open to multiple interpretation than other disciplines. Certainly the

vast majority of decisions as to acceptance or rejection are very straightforward for chemistry articles using traditional peer review.

The observation that different disciplines operate under so different conditions that entirely different quality control schemes may be appropriate should of course be taken seriously. However, with respect to the time delay to "publication," the question must be whether the chemists in this case want a quick decision in order to be able to disseminate the result to peers and obtain priority for it, or if it is in order to be able to put this additional merit item into his or her CV. If the former is the case then of course the delay time in the ETAI model is zero, since the result is disseminated and priority is established at the point where the review discussion starts. In the latter case, on the other hand, one will not be willing to accept substantial discussion periods, in particular if the character of the field is such that there is rarely much to discuss anyway.

In summary, we did have to work with explaining the two-stage open peer review model, and the important message had to be: in this system all the rules of the game are changed and all the habits change; you must think of it as an entirely different publication culture.

4.2. STARTING THE FLOW OF SUBMISSIONS AND DEBATE

Another type of problem involved starting the entire process: not only getting the first submissions, but also getting the discussion to start for each of these. This was a chicken-and-egg situation: people were not likely to contribute to a discussion that no one listened to, but people would only listen if there were already some contributions.

The relatively unsuccessful experiment with community peer review in *Nature* in year 2006 (Editorial Report: Nature's Peer Review Trial, 2006) may possibly be due to this problem.

Under the ETAI system, the interested community for an article was notified using an email message when the article was presented for review discussion. This was maybe sufficient for getting some of these researchers to take a look at the article, but it did not suffice for getting the discussion started.

Two measures were instrumental for dealing with this problem in the ETAI. When the journal was entirely new, we presented its review scheme as having some of the features of a conference presentation, besides being a journal. At a conference you can present your work and get feedback on it, but in our journal you could have 3 months of discussion instead of 5 min, and the discussion was open to everyone in the research community in question and not merely those that attended the conference, and finally it was preserved and could be read (and continued) later on. As a continuation of the same parabole we started panel discussions in the ETAI, where a few panelists made initial statements and then a discussion followed in our medium. This was effective in demonstrating to our constituency that if you send in a debate contribution then it is immediately seen by others, and this in turn encouraged submissions and debate contributions.

A second measure was taken if the discussion about a particular article did not start spontaneously: in those cases we could ask one or two colleagues to be discussion starters by making some initial comments. The experience was that once the discussion had started it tended to continue.

---

[3]www.aaas.org/spp/sfrl/projects/epub/ses3/parker2.htm

4.3. MAINTAINING COHERENCE

Our peer review model depended strongly on having an identifiable community whose members were likely to participate in the discussions. This was made possible by the fact that was mentioned initially, namely, that the research field of Artificial Intelligence is structured as a set of "virtual colleges" each having one or a few hundred members internationally. The mailing lists for the participants in these colleges were therefore essential for the functioning of the journal. Please recall that this was done long before the existence of social media; all communication had to be done using the journal's website and communication by email.

The ETAI was therefore organized as a federation of specific research areas, each with its own area editor, its membership list, and so forth. Articles could only be submitted to a specific ETAI area and if there was no area that matched a particular article then it simply could not be submitted. Area editors were quite independent and operated their own wings of the journal.

The coherence and uniformity of the journal therefore became an issue. In retrospect I feel that I should have done more toward building the team spirit in the group of area editors; this would have made the journal stronger, it could have resulted in amore uniform appearance in the websites of the respective areas, and most importantly, it could have given help and support to the area editors in their work.

At the same time I do not think it would have been possible to work without the organization as a federation of areas. The task of the area editor in this scheme requires expertise and recognized standing in the area in question. It also demands much more work than being an area editor in a conventional journal, in particular because the area editor has to moderate the discussions about the submitted articles.

4.4. INSUFFICIENT COMPUTATIONAL AND ADMINISTRATIVE INFRASTRUCTURE

The publication and peer review scheme that was used by the ETAI required a computational infrastructure for the following purposes:

- For the publication of submitted articles, using the Linköping University Electronic Press.
- For the dissemination of information about newly submitted articles, and for the reception and dissemination of contributions to the discussion about an article. This was done using both email messages to the area members and additions to the area's website.
- For the preparation of finally accepted articles in a form whereby the successive issues of the journal would have a graphic appearance that matched traditional journals.
- For the presentation of issues and volumes of the journal, containing both the actual articles and the review discussion for each of them.

These computational facilities were not ready when we started the journal; they had to be built as we went along. It would of course have been better to implement them first, but we had been eager to get started, we certainly underestimated the amount of work that was needed, and we did not know in advance what facilities

would be required. In any case, the requirement to build this software and, at the same time, to do the editorial work using partly improvised facilities led to a certain exhaustion on my part, and it was probably one of the factors that led to the discontinuation of the journal after a few years of relatively successful existence.

5. COMPARISON WITH CONVENTIONAL PEER REVIEW

An analysis of the advantages and disadvantages of a particular model for peer review should start with an identification of the goals that this process shall serve. Some such goals were mentioned in the Introduction, but there are in fact some additional goals that may be considered, as included in the following list.

- *Availability of reviewers:* insure that qualified reviewers will agree to participate and that they will wish to spend enough time and effort on the review assignment.
- *Amelioration:* improve the quality of a submitted article by providing feedback to the author.
- *Posterior use of reviews:* are the reviews valuable after the end of the peer review period?
- *Selection:* acceptance to the journal confirms that the article meets a specific quality standard, which helps readers decide which articles to read.
- *Fairness:* it is not merely in the interest of the readers, but also in the interest of the authors that acceptance decisions are fair and unbiased.
- *Merit:* acceptance of the article contributes to the author's scientific credentials.
- *Attention:* in the case of open-process peer review, the discussion in that process gives attention to the article in the researcher community of the article's topic.

We shall use this list as a framework for comparing the ETAI model for two-stage peer review with the conventional, blind review model.

The Attention aspect is by definition not present for conventional peer review. Authors in the ETAI reported that for them it was an important and positive aspect of the review model.

Conventional peer review integrates the Amelioration and Selection aspects into one single process. In two-stage peer review the two stages are dedicated to the Amelioration goal and the Selection goal, respectively.

The quality of the process with respect to Amelioration and Selection depends of course entirely on the competence and the efforts of the reviewers. I can only provide a subjective and qualitative estimate of this, based on also having been co-Editor-in-Chief of the journal Artificial Intelligence, AIJ (the most prestigious journal in its area) for a number of years, besides of course my general experience of other journals. My experience is that the quality of reviews varies enormously between journals, and that the quality of reviews (i.e., contributions to the open review discussion) in the ETAI was in the upper-middle range. It could not match the AIJ, but it was as good or better than many others.

One way of estimating the Selection effect is to check the acceptance rate, with an assumption that a low acceptance rate in a journal indicates that only articles with very high quality will be accepted there. In the case of the ETAI, the number of declined

articles was quite low. This might be an indication to its disadvantage, but there are some considerations that should also be taken into account. First of all, the numbers may not be comparable due to the "shame" effect that was discussed above: it is likely that authors thought carefully before submitting an article, in consideration of the risk of having it declined, and if this is true then the overall quality of submitted articles would tend to be higher. I have no way of quantifying this, but the argument suggests that one should be careful when comparing acceptance rates for the two peer review systems.

Another question in this context is whether it is truly in the interest of the scientific community that a journal is very restrictive with acceptances? For example, if reviewers have widely different assessments of an article and neither reviewer is willing to change their opinion, is this then a reason for accepting the article or for rejecting it? A strong emphasis on "quality" implies a reject decision, but this may effectively stop new and truly important contributions.

The usual argument in favor of a strict acceptance policy refers to the Selection goal: readers have limited time at their disposal, and the peer review process shall assist them by filtering out the articles that are required reading. Notice, however, that this is one more example of how the analysis departs from the characteristics of the conventional peer review system, without taking the effects of the alternative system into account. This is because in the conventional system, the *only* information that is available to the reader for his or her selection decision is the binary information that the article was accepted, plus of course the information about and by the author, such as the abstract. In the open-process model, on the other hand, the would be reader may check the *discussion* about the article as a first indication of whether the article is worth reading or not for him.

In general, the more metainformation you have about an article, the better. The abstract and the record of the discussion play different and complementary roles. As a reader, the information about the author and the author's institution gives some cues about quality and relevance. The title and the abstract are important for identifying whether the topic is relevant for him. The record of the discussion moderates these first impressions with respect to both quality, relevance and novelty. Consequently, a journal with open-process peer review may be somewhat more generous with its acceptances, thereby reducing the risk of missing important original developments, and still provide its readers with enough information that they can select their reading menu efficiently.

Another argument with respect to acceptance policies is that the acceptance of a marginal article tends to reduce the journal's impact factor. The argument goes as follows. It is known that the distribution of citation counts is extremely skewed, so that a small number of articles obtain very many citations, and most articles obtain few. Since the impact factor for a journal is calculated as the arithmetic average of the citation counts for all articles in the journal, any article whose citation count is lower than the journal's average will reduce its impact factor. Moreover, although one must be sympathetic to the problems of getting groundbreaking articles published, the hard fact is that they will only gain attention after a number of years, whereas impact factors are calculated based on citation counts during only a few years after publication.

Therefore, publication of such (rare) articles does not contribute favorably to the journal's impact factor.

The only thing one can say about this argument is that it illustrates the irrational character of the use of impact factors, and its detrimental effects on the scientific publication system.

The goal of Fairness is an important one. Benos et al. (2007) expressed doubt that open-process peer review would represent an improvement in this respect; they wrote:

> Both of these journals (ACP and ETAI) do not unmask the people who decide whether or not a paper is publication worthy. . . . This does not remove any bias, perceived or real, by referees or editors. Thus, these forms of open review, while alleviating the delays and increasing transparency, will not attenuate perceptions of bias at the actual acceptance step of the process.

This analysis is incorrect, for two reasons. First, the transparency of the review discussion and the attention that it provides for the article before the acceptance decision is a significant safeguard against malpractice in the refereeing stage. Secondly, even if an article is declined in the refereeing stage in the ETAI, it will still have the advantage of first publication with the ensuing citability and the proof of priority of the results. This means that a mistaken decision to decline or reject an article, should it occur, is much less detrimental for the authors and the article than what it is when the conventional peer review process is used.

A final remark concerns the Merit aspect of the peer review process. One consequence of the rapid growth of science and of scientific publication is that researchers and research projects are increasingly evaluated based on numbers that represent their publication and citation scores, whereas in older times it was taken for granted that in order to evaluate a person's research you must read and evaluate his or her publications. There are many voices to the effect that the numerical evaluation is very unsatisfactory, but the argument is anyway that we do not have any choice, in view of not only the amount of reading that would otherwise be required, but also the increasing specialization whereby reviewers are frequently called on to assess and to compare candidates whose area of research they do not themselves master. The persistent availability of the review discussion for an article may alleviate this problem, since even an outsider may often get a good notion of a researcher's standing and the quality of her work by hearing or reading an exchange of opinion between this person and his or her peers.

This possibility requires however that the discussion about each article is sufficiently extensive, which again adds to the reasons why it is in the interest of an author to have as many contributions to his review discussion as possible, including in particular a number of critical contributions that it is a challenge to answer.

6. MAINTAINING LIVELINESS IN PEER REVIEW DISCUSSIONS

As one can see from the ETAI webpage, some parts of the journal enjoyed lively peer review discussions, and in other parts the discussion did not really get off the ground. As stated in the Introduction, it is of great interest to understand the factors behind this difference.

6.1. PAST EXPERIENCE

Almost the first things that we learnt after starting the ETAI was that discussions do not usually start by themselves. Merely posting articles on the journal's website and inviting contributions is not very effective. I have described the methods that we used for starting discussions, and some of the cases of failed discussions may have been due to the insufficient use of these methods.

However, looking in retrospect at the ETAI experience it seems that another factor was also important, namely the question of *reader fatigue* and the related question of *limited exposure*. In those cases where a reader of the journal was exposed with a considerable number of articles in the same short period of time, it seemed that it was difficult to get the reader to engage herself or himself in any of these articles, whereas if only a few articles were offered and these were quite relevant to his interests, then it was much more likely that he or she would write a debate contribution. The partitioning of the journal and the readership into areas of limited size insured that each reader of the journal received a sufficiently limited exposure and a sufficiently focused set of new articles per time unit for her or his consideration.

The hypothesis that a limited reader exposure was important for insuring good participation in the discussions is not something that we can validate by hard data; it is only based on a general understanding of how our readers operated. It is however consistent with the actual discussion intensity in the ETAI, and in particular with the outcome of our attempts to base special ETAI "sections" on contributions at specialized workshops. The idea for this was simple: such a workshop engages the same "virtual college" as is used for defining an Area within the ETAI, workshops are used both for presentation of recent work and for discussions, and the ETAI seemed to be a natural way of extending both those aspects of the workshop activity. To begin with, we would invite the workshop participants to write down their main comments at the workshop and to contribute them to the ETAI.

This worked very well in one case, and not very well in several others. The *Special Section on Knowledge and Reasoning in Practical Dialog Systems*[4] is a case where it worked very well, but it also required a considerable effort by the area editors for obtaining and editing the debate contributions from the workshop participants. On the other hand, when individual articles were submitted one at a time it was easier for an area editor to obtain a viable discussion.

It is interesting to compare this experience with the situation in the journal *Atmospheric Chemistry and Physics* (ACP; Koop and Pöschl, 2006)[5] which is arguably the most successful example of two-stage open-process peer review at present, and which started its operation in 2001. The peer review procedure in the ACP, as described on its website, is in principle quite similar to the one used by the ETAI, but with one major difference: the ETAI was organized as a federation of areas and the discussion was primarily viewed as an internal discussion within each area, but the ACP does not have such a structure. All submitted articles are presented in a single, chronological list on the ACP webpage, and the interested reader

will see all of them. Furthermore, the publication volume of the ACP is significantly higher than for the ETAI.

It is against this background that one must read the statistics about the participation in review discussions in the ACP. For example, as observed on May 15, 2011, among the 41 submissions that had been received between March 1 and March 15, 32 had obtained no or one contribution to the discussion. Six of them had obtained 2 contributions, and 3 of them had obtained 4 contributions. Among the 24 discussion contributions in the discussions with more than one contribution, only 5 where by third-party persons and the other 19 were by a designated referee or by the authors. These figures apply 2 months or more after the beginning of the discussion. For the 39 articles received between May 1 and May 15, only one of them had even one discussion contribution.

It seems, therefore, that although the ACP is a very impressive example of the use of open-process peer review, the most important aspect of its model is that it advances the transparency of the review process, and that it guarantees that articles are published and citable from the very beginning of that process. On the other hand, if one is interested in obtaining a real community discussion about submitted articles, then the ACP does not offer a strong case.

As already mentioned, the approach used by the ETAI was relatively labor intensive for each of the area editors, and it only covered some parts of Artificial Intelligence. Consider, therefore, the question how one would organize a journal that used open-process peer review with lively discussions and that was anyway able to publish several hundred articles per year. How would it be organized, given what has been said about the need to both encourage and to moderate the discussions about each article. This is the question that must be answered if the strong aspects of the ETAI experiment is going to scale up.

6.2. A FIRST PROPOSAL

The first step toward answering this question must be to obtain a clear understanding of the structure of the scientific discipline that the journal would serve. Does it resemble the structure of Artificial Intelligence where there are identifiable specialities with their own problem statements, memberships, workshops, cooperations, and competitions, and is the difference only that the number of such specialities is much larger? Alternatively, does it instead have a more open structure where researchers continuously monitor research articles and results that emanate from a much larger population of fellow scientists?

In the former case I imagine that it should be possible to scale up the approach that was used by the ETAI while using the Wikipedia organization as a model. Concretely speaking, it would be necessary to organize the resulting large number of areas and area editors using a firm set of rules and guidelines for all aspects of the journal's operation, and to have a reliable and complete computational infrastructure already from the start of the operation. These were things that the ETAI did not have.

6.3. A SECOND PROPOSAL

In the latter case, it seems clear that the ETAI model would not work: having a large number of members in an area for the journal

---

[4]http://www.ida.liu.se/ext/epa/ej/etai/1999/D/
[5]http://www.atmos-chem-phys.net/

would put an unreasonable workload on the area editor, and our informal observations of the importance of reader fatigue suggests that participation in the discussions would anyway be too low. Moreover, the observations of actual debate participation in the ACP suggests that its model will also not be able to support lively discussions.

I will therefore offer the following proposal for how to organize a larger journal in this case: one may try using a system based on *ad-hoc discussion groups.* For each article, or for a small set of related articles, one would form a discussion group that should last for the entire review period of the article(s) in question. Peers should not be enabled to make discussion contributions randomly in the full set of articles that are under discussion, but only by joining a discussion group and staying with it. In order to insure continuity and coherence in the discussion, a participant in the journal's discussion activities could be encouraged to engage in a reasonable number of groups at each point in time, and to join a new group when one that she is in has completed its work, i.e., the acceptance decision has been made. The identification of a new group to engage in could be made through invitation by another group member ("Here's an article that you'd find interesting") or by active search by the participant, or by a service where the software system suggested relevant groups.

An important consideration would then be to strive for a good mix of participants in each ad-hoc group, in particular, to engage the entire range from Ph.D. students to senior researchers. In fact, an advisor might find it worthwhile to *require* her or his students to participate in a number of such groups as one part of their Ph.D. study.

The purpose of organizing such ad-hoc discussion groups would be to arrange a level of contact between reader and journal where limited and focused reader exposure is obtained, and where it should be possible to attract and retain the reader's attention to a limited number of articles. An obvious problem with this model would be that some articles may attract a very large number of discussants, and others may not attract any. The former problem should not be handled by creating several groups, since it would overburden the author; it would be better to simply let the system enforce a limit on the number of discussants for each article. The problem of no discussants or too few discussants is more difficult, but one possibility would be to refer such articles to conventional peer review.

Another possibility would be to decide that if no one is interested then the article is automatically declined for the journal. Such a policy would not be as harsh as it may sound, since the likely of effect of it would be that each author would try to engage a certain number of discussants for her or his article. Hopefully this would be sufficient for avoiding the situation where a perfect paper is dropped because no one has anything critical to say about it. The scheme might however bias the discussion in a too positive and uncritical direction. This can only be determined by actually experimenting with this policy as well as alternative ones.

7. ADDITIONAL ASPECTS OF TWO-STAGE PEER REVIEW

Although the question of maintaining liveliness of discussion even in the case of scaling up is the most important issue, there are anyway some other aspects of two-stage open-process peer review

that may be discussed in the light of the experiences that have been described.

7.1. SHOULD OPEN-PROCESS PEER REVIEW USE AN OPEN-NAMES POLICY?

With the experience from having operated the ETAI it is interesting to read about other experiments with open-process peer review as well as reading more general comments and proposals in the same direction. It is striking that many of them make the same extrapolations from the culture of conventional peer review as we encountered when the ETAI was started. In particular, it is frequently argued that the identity of the discussants must be kept confidential because otherwise the comments will be very dull; see e.g., Suls and Martin, 2009), or Khan (2010) for an editorial in the British Medical Journal. Our experience was however contrary to observations such as these, for the reasons that were stated above.

There was in fact one particular occasion when a discussant requested that his name should be withheld, but for an interesting reason: he had made similar, critical remarks to the same article when it had previously been submitted to a conventional journal, and rejected, and if his name were to be stated in the ETAI discussion then he feared that the author would be tired because of the role he had played in the decision of that other journal. This illustrates how it is the character of the conventional peer review process that *causes* reviewer anonymity to be an issue, and not the phenomenon of critique in itself.

To the extent that lively review discussion is considered as an important goal, so that transparency of the review process is not the only consideration, it is also plausible that an open-names policy with respect to all participants in the discussion will increase the attention that is paid to the discussion, and therefore, will tend to increase the number of further contributions to it. Knowing who has written a contribution to a discussion adds to the reader's perspective on it and is likely to stimulate her or his opinions on the matter. It follows also that an additional advantage of the open-names policy is that it may help strengthening the community of researchers in question, and in particular to help including those that are not able to travel to the important conferences.

7.2. DURATION OF THE COMMENTARY PERIOD

Several proposals for open peer review suggest that the discussion should go on for an unlimited time, and in some cases that there should not be any strict acceptance decision but merely an initial screening for relevance and appropriateness of a submitted article. This means in effect that only the first stage of the ETAI two-stage process is used, and it goes on indefinitely. However, even in the two-stage process there is absolutely no reason why one should not be able to add further comments to the discussion after an article has been accepted, or after it has been declined, and in the latter case this might also lead to the article being reconsidered for acceptance[6]. On the other hand I still believe that there is a value in having a limited period of time when particular attention is given

---

[6]This indicates in fact an additional advantage of open-process peer review: if an article has been declined mistakenly then the mistake can be corrected later on and the author can receive due credit. In the conventional peer review system it is very difficult to correct such mistakes.

to the article, so that one can obtain a coherent discussion about it and not merely a number of occasional comments.

The question of what is the optimal duration of the commentary period is an important one. If it is too short then it will not give peers enough time to think and to react; if it is too long then peers may be led to postpone making their contributions, which leads to a loss of dynamism in the discussion. Moreover, the observation concerning reader fatigue suggests that commentary periods should be kept short, so that the set of articles under discussion at any one time is kept fairly small. Different journals and different disciplines may strike this balance in different ways. In the case of the ETAI I think the 3-month period was reasonable, but 2 months would probably also have worked well.

7.3. ARTICLE PUBLICATION STATUS DURING THE REVIEW PHASE

An additional difference between the peer review procedures in the ETAI and the ACP concerns the publication of articles at the beginning of the review debate. In the design of the ETAI procedure we were very concerned about the publication status of a submitted article during its discussion period, and as explained above we defined a mechanism whereby the article would count as *published* on the date when it was advertised and made available to its peer community for the purpose of discussion, in particular so that it would count for priority of results. We created the Linköping University Electronic Press for this purpose, and we participated in the discussion at that time about what constitutes an electronic "publication."

The ACP has chosen another approach: concurrently with the ACP journal there is the journal-like *Atmospheric Chemistry and Physics Discussions* (ACPD) whose webpage is graphically similar to its parent journal, but where it is made clear that articles are included there prior to peer review and eventual acceptance in the ACP.

The approach used by the ETAI was more elaborate. We chose it because of a long-term consideration where we wanted research articles to be associated with research data and with computational processes that illustrate and validate the contents of the articles themselves. Such attachments to articles impose particular demands with respect to long-term maintenance, and it was not possible to make such guarantees in our E-Press for all ETAI authors that might wish to use such facilities. Instead, the strategy was to encourage other institutions in our area to set up their own counterparts of the E-Press, so that both the pre-review publication of the article itself and the definite publication of the attached resources should be done in the author's home institution, or in an entity that was dedicated to this service – a kind of "web hotel" for research articles and their related materials.

It turned out that no other institution reacted to this suggestion during ETAI's active period, so in practice the Linköping E-Press ended up doing the initial publication of all submitted articles, as well as of course the ETAI journal itself. However, I still believe that the proper organization of attached computational materials is an important issue for the future, at least for our field of research and probably for many others.

Another consideration with respect to publication status and priority arises with respect to how we defined the date of publication of an article. Since we considered in principle that the starting date of the discussion period was the date of publication of the result, we used it for defining the date of publication of the final article. Thus an article whose discussion started in October of year X and that was accepted for the journal in February of year $X + 1$ would appear in the journal issue for October-December of year X. The logic behind this was clear, but it was not always easy to explain it to authors and readers.

This design led in turn to another consideration, namely, a restriction on what changes were permitted in an article between the original submission and the final version for the journal. On one hand we wished of course that the review discussion should result in improvements, but on the other hand it would have been unfair if the final version were to contain essential results that had been obtained after the publication (in our sense) of the first version. There was a rule, therefore, that the changes should be restricted to improvement of presentation, without strengthening the results as such.

In one concrete case, an author of a relatively theoretical article reported during the discussion period that he had some additional results that would fit well into the same article, and the question was what to do with them. The solution was that his additional results were written up as "short note" that was presented as an addition to the original article, but with a later date of publication. Such a separation of the results would have been inconvenient in a paper-based journal, but in the electronic medium it was not a big issue.

These considerations with respect to publication date may seem unnecessary, but my view on this is that they should be viewed in the same way as formal business contracts in one's personal life: as long as the relations between people are dominated by common sense there is no need for formality, but if problems should arise then they can be handled with less pain if there are clear rules and clear data. Priority of research results is sometimes a topic of considerably animosity, and it is worthwhile to design one's publication system in such a way that one has a firm basis for resolving conflicts at those rare occasions when they do arise.

7.4. INNOVATIVE SOFTWARE TECHNIQUES VS. CLASSICAL STYLE

Several of the measures that we took in order to make the ETAI acceptable are no longer needed, and may be irrelevant for future introduction of two-stage peer review. We organized our journal in terms of annual volumes and issues, with consecutive page numbering throughout each volume, although in principle it would have been more natural to consider an annual volume just as a set of articles and to number the pages of each article from one and up. We also produced a small supply of paper-printed copies of each issue, with a nice-looking cover, so that we could show it at conferences and archive it in major libraries. Measures such as these are superfluous today, or will soon be.

The computational infrastructure that was used by the ETAI seems antiquated by contemporary standards. Today we would certainly use a more interactive implementation. It would be natural to consider using wiki techniques and social-media techniques.

At the same time I would be careful not to go overboard with the use of modern software paradigms. For good and for bad, prestige is an important factor for a scientific journal, which means it must

inspire confidence and signal continuity. This applies not only for the articles that are submitted, debated and eventually accepted, but it applies as well for the discussion. In the case of the ETAI we made sure that the discussion contributions were presented in a correct fashion. In fact, one of the ETAI areas actually operated a side-journal called an Electronic Newsletter that was dedicated to presenting the discussion contributions, as well as other information of interest, in a nicely formatted form that resembled the format of the main journal. This was done in order to give prestige, in a good sense, to the discussion contributions so that people should feel that these discussions were valuable material: valuable to read, and valuable to have written, something that you could add to your C.V.

One other aspect of the prestige policy was to maintain a high conversational standard in the review discussions, besides of course a high scientific standard. The discussion was moderated, no contribution appeared on the website until it had been approved by the area editor, and the tone of critical comments was monitored. In fact, it is not so uncommon that reviewers in conventional peer review take advantage of their anonymity for adopting a condescending tone vis-a-vis the author and the submitted article. Some discussants retained the same haughty attitude in their contributions to our discussion. We therefore imposed a strict policy of asking the discussant in such cases to revise the wording and to adopt a tone that he would use if he talked to the author face to face and in a civil manner.

My suggestion for a contemporary open-process peer review scheme would therefore be to carefully consider all that can be offered by modern Internet-related technology, but to only adopt it when it is compatible with a policy of consistently good style and effective quality control of all aspects of the journal's operation.

7.5. BEYOND CONVENTIONAL ARTICLES: PEER REVIEW IN NEW ENVIRONMENTS

Innovation in the publication and communication of research results is not confined to the well-known topics of electronic publishing and open access, or to the current topic of changing the peer review model. The present article has discussed alternative peer review but with an assumption that the character of the articles themselves has not changed. This assumption will not remain valid for long. There is an abundance of new topics when other kinds of publications are considered, and here I can merely indicate my own particular interests in this respect. One important topic concerns the organization of *evolving articles* where the author of an accepted article is made responsible for the update and maintenance of the article during a period of time and is able to amend it successively (Sandewall, 2010). I am also interested in the question of publication of information modules whose contents range from "facts" to "knowledge," and how such modules can be published, peer reviewed, cited, and so forth (Sandewall, 2008, see also the Common Knowledge Library[7]). Finally there is an interesting issue concerning how to organize a publication mechanism that is appropriate for publishing the results of large, integrated, systems-oriented projects. All these new kinds of publications will require novel forms of peer review that are adapted to

their peculiar characteristics. I am convinced that an open-process peer review scheme will be appropriate in those cases as well, but the basic setup will be different from what you need for peer review of conventional articles.

7.6. COEXISTENCE BETWEEN PEER REVIEW SCHEMES

One of the most important observations from the ETAI experiment is that open-process peer review creates and requires a culture that differs from conventional peer review in important ways. The change of rules and practices affects the expectations and the behaviors of authors and of reviewers in ways whereby these behaviors tend to gravitate to a new and different equilibrium, so to say.

This raises the question as to what will happen when conventional and alternative methods of peer review coexist. Several scenarios are possible. One may imagine a polarization where some research communities embrace the new methods wholeheartedly and other communities reject them outright. One may also imagine the emergence of intermediate models: a kind of "open peer review light." Finally one may imagine a kind of "survival of the fittest" in the competitive world of research publication, namely, if the disadvantages of belonging to the minority that uses a non-standard scheme are so big that it can not survive in the long run. For example, quantitative research assessment constructs such as impact factors and acceptance rates are based in the culture of conventional peer review, and furthermore they tend to favor existing journals over new ones. If they are applied to publication venues that use alternative peer review schemes then these may easily find themselves at a disadvantage in several ways.

8. CONCLUSION

In this article I have discussed the experience from the Electronic Transactions on Artificial Intelligence and made some suggestions for what would be needed in order to scale up the size of a journal with open-process peer review without sacrificing the liveliness of the review discussion. An additional theme of the article has been that the use of the combination of open-names and open-process, two-stage peer review tends to change the researchers' perceptions and expectations in the review process in a multitude of ways, and that it can easily be very misleading to try to predict what will happen in such a scheme by extrapolation from what is the case when conventional peer review is used.

This observation is in opposition to a suggestion made by Stevan Harnad when he wrote as follows (Harnad, 1997):

> Peer review is imperfect; it can no doubt be improved upon, but alternatives should first be tested; and in testing, one is well-advised to manipulate one variable at a time: Here we are dealing with a change in medium (paper to electronic), a change in economic model (subscription to author-side payment) and a change in quality control mechanism (peer review to open peer commentary).

As we have seen there is a number of other "variables" that are also being changed, and the problem is that the effects of those changes are not independent. There are clear indications that when a change of one variable at a time is likely to have one set of consequences, the effects of changing several of them together may have

---

[7]http://piex.publ.kth.se/ckl/

consequences that are quite different from the individual changes. This is a reason why the topic of alternative methods for peer review is so difficult to analyze, and such a fascinating challenge to experiment with.

REFERENCES

Angell, M., and Kassirer, J. P. (1991). The Ingelfinger rule revisited. *N. Engl. J. Med.* 325, 1371–1373.

Benos, D. J., Bashari, E., Chaves, J. M., Gaggar, A., Kapoor, N., LaFrance, M., Mans, R., Mayhew, D., McGowan, S., Polter, A., Qadri, Y., Sarfare, S., Schultz, K., Splittgerber, R., Stephenson, J., Tower, C., Walton, R. G., and Zotov, A. (2007). The ups and downs of peer review. *Adv. Physiol. Educ.* 31, 145–152.

Editorial: Revolutionizing Peer Review? (2005). *Nat. Neurosci.* 8, 397.

Editorial Report: Nature's Peer Review Trial (2006). *Nature.* doi: 10.1038/nature05535

Frankel, M. S., Elliott, R., Blume, M., Bourgois, J.-M., Hugenholtz, B., Lindquist, M. G., Morris, S., and Sandewall, E. (2000). Defining and certifying electronic publication in science. *Learn. Publ.* 13, 251–258.

Gura, T. (2002). Scientific publishing: peer review, unmasked. *Nature* 416, 258–260.

Harnad, S. (1997). *Listserv Comment on 'Open Peer Commentary: A Supplement, Not a Substitute, for Peer Review.'* Available at: http://list.uvm.edu/cgi-bin/wa?A2=ind9706&L=serialst&D=0&P=4500&F=P

Hodkinson, M. (2007). Open peer review and community peer review. *Journalogy.* Available at: http://journalology.blogspot.com/2007/06/open-peer-review-community-peer-review.html

Khan, K. (2010). Is open peer review the fairest system? No. *Br. Med. J.* 341, c6425.

Koop, T., and Pöschl, U. (2006). An open, two-stage peer review journal. *Nature.* doi:10.1038/nature04988

Parker, R. (1998). "Response to Sandewall's alternative view to peer review," in *AAAS/UNESCO/ICSU Workshop on Developing Practices and Standards for Electronic Publishing in Science*, Paris.

Sandewall, E. (1997a). A neo-classical structure for scientific publication and reviewing. *Linköping Electronic Articles on Academic Policies and Trends.* 2, nr 1.

Sandewall, E. (1997b). Publishing and reviewing in the ETAI. *ETAI* 1, 1–12.

Sandewall, E. (2006). Opening of the process. A hybrid system of peer review. *Nature.* doi:10.1038/nature04994

Sandewall, E. (2008). Extending the concept of publication: factbases and knowledgebases. *Learn. Publ.* 2, 123–131.

Sandewall, E. (2009). "Experience of two-stage peer review in the ETAI, 1997-2001," in *International Symposium on Peer Review*, Orlando.

Sandewall, E. (2010). Exercising moral copyrights for evolving publications. *ScieCom Info* 6, 1–4.

Suls, J., and Martin, R. (2009). The air we breathe. A critical look at the practices and alternatives in the peer-review process. *Perspect. Psychol. Sci.* 4, 40–50.