How to improve the quality of peer reviews? – Three suggestions for system-level changes

Abstract

Peer reviewing is critical in the process of legitimizing new scientific knowledge. Yet, there are concerns about its quality, especially if one considers developmental reviewing as an ideal. The present essay suggests three ways to improve review quality: provide reviewers with systematic feedback about their performance, reward active and good reviewers, and make reviewers more accountable by revealing their identities to the authors on certain conditions.

Introduction

Scholarly or scientific peer review is the evaluation of research findings for competence, significance and originality, by qualified experts who do research in the same field (Brown 2004, Benos et al. 2007). It is critical in the process of legitimizing new scientific knowledge and assuring its quality. A piece of research that has not passed scholarly review and has not been published cannot be regarded as scientific, since its findings have not been accepted by the scientific community in question and are not trustworthy in that sense.

Just as democracy peer reviewing it is not perfect, but nobody has found a better alternative to it. Benos et al. (2007) summarize its weaknesses such as various biases (status and gender biases, and biases because of ideological differences, unconventional ideas and conflicts of interest), its inability to identify major flaws and scientific misconduct, and delays in the publication process. Yet, scholarly reviewing provides an opportunity to authors to respond to the points of criticism raised by peers before publishing and consequently to improve the articles. That alone is a sufficient reason to preserve peer reviewing (Benos et al. 2007).

In the field Information Systems (IS) not only the quality of peer reviews is a concern (Gray et al. 2006), but there are also opinions that the quality of reviews has recently been deteriorating, as evidenced with the discussion on the AISWorld forum during Fall 2013. One reason is the increasing number of journal submissions, leading to a constant shortage of (good) reviewers. If the quality of reviews is getting lower, there is a higher risk that journal editors make Type I errors, in which papers of low quality are accepted, or still more alarmingly Type II errors, in which papers with great potential are rejected (Straub 2008).
Therefore scientific communities should always consider how peer review can be improved. But what is a good review? Carpenter (2009) proposes that a great review “identifies weaknesses and identifies a path or paths to remedy those weaknesses” (p. 193), i.e. they are developmental (Saunders 2005a, Saunders 2005b, Lepak 2009). A developmental review is not only of high quality in the substance of the manuscript - in its evaluation of the contribution, theoretical background, research method, argumentation, and presentation – but it is also constructive in all its critique, friendly expressed, supportive, detailed enough, includes concrete suggestions, and overall is helpful when improving the paper. It is not written to the editors only but also - and equally importantly - to the authors of the submission. In the present essay I regard a developmental review as an ideal.

There are some published guidelines for reviewers, how to write good reviews (such as Lee 1995). There are also some discussions of what developmental review means in particular from the viewpoint of reviewers (Saunders 2005a, Saunders 2005b, Lepak 2009). Although definitely useful, they rely on individual reviewers’ readiness to develop their reviewing skills and do not imply any system-level changes that might create conditions for better reviews and motivate reviewers to invest more on reviewing.

There are also some suggestions on how to use IT to support the review process (e.g. Mandavwialla et al. 2008, Kane et al. 2009, Hardaway and Scamell 2012). However, when resorting to the technology one should consider the rules of conduct to be embedded in the technology and the larger system of publishing.

The present essay suggests three technology-independent system-level means to improve the quality of reviewing – provide systematic feedback to reviewers, reward for good reviews, and make reviewers more accountable by revealing reviewers’ identity to the authors on certain conditions. Before proceeding to these proposals, as a motivation I will introduce some of my personal experiences of reviews.

**My personal experiences as a reviewer and as an author**

Since most of the reviewing is voluntary work that is not directly rewarded in any way, it may be next to arrogant to criticize the quality of reviewing. Especially journal editors often must be happy if they manage to recruit a sufficient number of reasonably qualified reviewers, who are ready to do the job. Yet, one should always keep in mind that reviewers act as significant gatekeepers in science. Their reports essentially affect not only editorial decisions on the submitted articles, but sometimes the future of the authors, and in the longer run the direction of the journal will take, and the progress of research of the disciplinary field as a whole. Therefore, research communities cannot omit the quality issue of reviewing just for reasons of discreetness, but it should also be subjected to a critical scrutiny.

I have not kept any record of my reviewing appointments, but I also guess that I have reviewed a few hundreds journal manuscripts during my academic life. I have also served in editorial positions in journals such as *European Journal of Information Systems* (2003-2007), *Information Systems Journal* (1997-2012), *Journal of the Association for Information Systems* (2005-2008), and *MIS Quarterly* (2007-2011) and in those positions handled the review processes of quite many manuscripts. Furthermore,
I estimate that I have personal experience of 60-70 journal review processes as an author.

As can be expected, the quality of the reviews I have seen is close to the normal distribution. Most of them have been average, not particularly good, not particularly bad but reasonable, some have been excellent, and some weak. It is hard to change this distribution, since it is a kind of human invariant, typical to the quality of any human activity. Yet, journals could attempt to move the peak of the distribution towards better reviews and reduce its variance especially by eliminating bad reviews.

Bad reviews are reviews that are substantively of low quality, indicate excessive bias in its evaluation, exhibit a negative (non-constructive, non-developmental) attitude towards the manuscript, and/or do not have any respect for the author.

One reason for bad reviews is the assignment error when a wrong reviewer is assigned to a paper. An assigned reviewer may not be knowledgeable enough about the topic of the paper or the research method applied in it, (s)he may be positively or negatively biased because of his/her paradigm and worldview, (s)he may have a conflict of interest that compromises his/her ability to review objectively, or simply (s)he does not have interest in the submission or sufficient time to review properly. It is a constant challenge of journal editors to avoid such assignment errors.

A second reason is more attitudinal. Some reviewers, possibly in some circumstances, are non-constructive rather than constructive in their reviews. According to my experience such cases are really rare, since I have personally encountered only two. The first one was when I submitted a paper to a premier IS journal in the late 1990’s. The paper attempted to understand IS development in terms of number of inherent dilemmas preliminarily discussed in Iivari (1996). I do not have the review report of this submission archived anymore, but the most memorable remark in it was the comment that the paper looks like it has been written after having had a few drinks. I did not find this remark particularly friendly or even funny.

The second case is more recent. I submitted a paper on two strategies of design science research (DSR) (Iivari 2012) to another top IS journal in 2011. The paper was inspired by the review process of Sein et al. (2011), in which I served as an associate editor. During this process I was forced to contemplate the relationship between Action Design Research (Sein et al. 2011) and DSR (Hevner et al. 2004). This gradually led me to distinguish the two DSR strategies. The first strategy corresponds to mainstream DSR, in which a researcher constructs or builds an IT artifact as a general solution concept to address a class of problem, and possibly instantiates it to address a specific problem. In the case of the second strategy a researcher attempts to solve a specific client’s problem by building a concrete IT artifact in that specific context and distills and generalizes from that experience prescriptive knowledge to be packaged into a general solution concept. I used Markus et al. (2002) and Sein et al. (2011) to exemplify this second DSR strategy.

The first review cycle of this manuscript took almost 1.5 years. One of the reviewers did not find any contribution in the distinction between the two DSR strategies and commented that “I really doubt that you are 'introducing' these choice, they have been described previously in papers that you cite, e.g., Markus et al, Sein et al, and Peffers.
et al”. According to my reading none of these references identifies and still less contrasts the two DSR strategies. So, I interpret that the reviewer suggests that any analysis of schools of thought in philosophy, sociology, economics and so on makes no contribution, since those schools have already been identified by researchers classified to represent those schools.

This reviewer continued “In the major part of the paper you use Markus et al and Sein et al as examples to contrast the two ‘strategies.’ This is a rather thin basis for you to make a contribution. What does the reader learn from reading your paper, rather than just those two papers - the only ones on which your analysis is based?” To me, this comment is equally friendly as feedback to a student after an exam that even though (s)he was able to write his/her name correctly it is not enough to pass the exam.

Furthermore, this same reviewer had difficulties to understand the distinction between DSR and behavioral science research, which is fairly standard in the DSR literature (e.g. March and Smith 1995, Hevner et al. 2004). (S)he questioned “why contrast DSR with behavioral research only? What about economic based research, interpretive research, other kinds of qualitative research, formal proofs, grounded theory, socio-technical, gender-based and any of a number of different paradigms used in IS research?” Yet, the associate editor (AE) handling my submission assured in his/her AE report that the question is about a seasoned expert. Assuming that the AE was right it seems to me that this reviewer had serious attitudinal problems with my submission.

In conclusion, I think that the two examples illustrate that there are really bad reviews, which are disrespectful, unfair and non-constructive. Even though rare, they may be catastrophic especially to junior authors. They may weaken their chances to be promoted, destroy their academic careers, and possibly ruin their lives. In the following I will suggest three system-level means to improve the reviewing process, means that hopefully are able to reduce weak reviews and to eliminate really bad ones, or at least that to prevent that they are continuously repeated by the same reviewers.

**Provide systematic feedback to reviewers**

I have never encountered a system in which a reviewer is provided with genuine feedback about his/her performance. It may be that some journals collect information about the turnaround time of reviews in the case of each reviewer for editorial purposes, but I have never received such feedback as a reviewer, i.e. how I have performed when compared with my peers.

Thorough reviewing takes its time, but it is much more crucial that reviews are substantively and attitudinally of good quality. At least I have found it difficult to evaluate the quality of my reviews. As in case of any piece of text, the author is not the best to assess it. Furthermore, it is not easy to know how the readers of the review, i.e. the authors of the manuscript and the editor, have perceived the review. Therefore it would be good to reviewers to receive systematic feedback about their reviews.
Currently, the only way to get some feedback about one’s review is to compare it with those of other reviewers and with the AE’s report\(^1\), if the whole review package is delivered to all reviewers as it is usually done. If your review is in line with those of other reviewers and if the AE picks up some of your points in his/her report, you are safe in a way.

Assuming that an ideal review process is developmental, it is significant to receive also the authors’ view, i.e. did the authors experience that the review was substantively of high quality, detailed enough, actionable, kindly expressed, and so on? There could be a short questionnaire to collect such information. The collection from authors could take place after each review cycle so that it would be as authentic as possible, but delivered to the reviewer – if permitted by the authors – only after the case is closed, i.e. when the paper is finally accepted or rejected so that it does not affect the ongoing review process. If preferred so, the feedback could be delivered to the reviewer in an aggregated form when (s)he has a few reviews completed (five, for example). In the case of rejected papers, the feedback can be anonymous.

When the paper is finally accepted or rejected, the AE can also make his/her own evaluation of each reviewer based on his/her own assessment, but taking into account the authors’ feedback. Especially if the AE’s assessment of a reviewer considerably differs from that of the authors, (s)he should justify the difference. It may well be that the AE evaluations of individual reviewers cannot be delivered at this stage, since it might weaken his/her opportunities to recruit reviewers in the future. However, once a reviewer has five or so reviews the feedback could be delivered to him/her in an aggregated form, especially if the reviewer has served several AEs.

The system described above would make it possible to identify good reviewers, whose reviews are consistently found to be of high quality and developmental, and also those reviewers whose reviews have consistently been non-developmental or otherwise of low quality. The former can be rewarded for their good performance. In the case of latter reviewers there is a good reason to critically look at their reviews and consider measures to correct the situation.

**Reward great reviewers**

Scholarly reviewing is voluntary work that is not directly paid for and compensated in any way. It is absolutely necessary for the scientific community to work, but according to my experience active and great reviewing is not much counted, when one’s academic performance is assessed. It is hard to imagine that this would change in the near future, even in the case of open access journals, although they may change the business logic of scholarly publishing.

Of course reviewing helps junior researches to understand the publication game and to enter the scientific community. A reviewer may also receive some valuable intellectual capital earlier than others, but (s)he must be very cautious with it so that no-

---

\(^1\) Although all journals do not make a distinction between senior editors and associate editors, for simplicity the term “associate editor” is used in this essay to refer to those editors who usually recruit reviewers and write the first editorial report based on those reviews and their own reading of the manuscript.
body cannot claim that (s)he has stolen an idea from some rejected manuscript, for example.

If a great reviewer advances to an editorial position, (s)he gets visibility, recognition and influence in his/her research community and may be rewarded in that way. Yet, there are relatively few who are “promoted” to such positions. Furthermore, since journals do not have systematic evaluation of reviewers’ performance, the selection of new editors is usually based on other merits such as the number of publications in prestigious journals and personal contacts rather than on great reviewing.

As a whole, reviewing is invisible work, conducted largely for altruistic reasons. At the same time most reviewers work under an increased pressure to publish. This leads to a dilemma between the altruistic motivation to serve the community and the egoistic motivation to publish more and to satisfy the expectations of his/her employer or even to exceed them. I am afraid of that the “publish or perish” push forces most potential reviewers to emphasize the egoistic desire to advance one’s career rather than the altruistic service to the research community.

In this situation it is becoming increasingly important that also good reviewing is concretely rewarded so that the egoistic motivation to publish and the altruistic desire to serve the community can be better aligned. One option for journals to concretely reward active and good reviewers is to introduce a new category of articles (“reviewer forum”), which provides great reviewers with a “special” opportunity to publish. Of course, this should be complemented with a fair system in choosing the great reviewers. To underline it the published reviewer forum article could also be supplemented with a summary of the author’s contribution as a reviewer in the journal.

These reviewer forum contributions do not necessarily compromise the scientific quality of the journal. They can be subjected to peer review just as other articles, even though with more modest quality expectations than full research articles. Most IS journals already publish editorials, research notes, opinion papers and so on with their own quality requirements. So, why not reviewer forum articles?

I suppose that the two system-level changes discussed above are fairly easy to adopt and implement, since I do not see that reviewers, editors or authors had a specific reason to resist them. The third proposal on the contrary is more controversial, although the reviewer anonymity has been questioned at least for 30 years (Armstrong 1982).

**Make a reviewer accountable by revealing his/her identity to the authors**

Everybody who has followed reader comments in the context of online newspapers, for example, has evidenced that the author anonymity is a significant reason for uncivilized and offensive comments: if the commentators are anonymous, the quality of comments may be quite low; when commentators use their real names, comments are much more carefully considered and more constructive. Scholarly research has also confirmed this observation (Santana 2014). Anonymity leads to lower self-control and lower quality of discourse (Ruesch and Krämer 2012).

Taking analogy, one can seriously question if the reviewer anonymity is the major reason for reviews of low quality and in particular of non-developmental reviews.
Yet, the double-blind reviewing, in which the reviewers do not know the authors and the authors do not know the reviewers, is often regarded as the preferred form of peer reviewing. The two-way anonymity is believed to increase objectivity of the process and reduce biases (Hillman and Rynes 2007).

But is it really so? Regehr and Bordage (2006) identify four alternatives to double-blind reviews: single-blind in which the authors are revealed but reviewers concealed, single-blind reversed in which the authors are concealed but reviewers revealed, optional single-blind in which the authors are revealed and reviewers are free to sign, and open review which is entirely open. This classification misses optional single-blind reversed review in which the authors are concealed but reviewers are free to sign.

Although the decision between single-blind review and double-blind review is not any big issue in top IS journals, it is interesting to note that the evidence supporting the latter is not so compelling. While Hillman and Rynes (2007) claim that existing research suggests that revealing the identity of the author influences the acceptance so that papers from prestigious universities are more easily accepted particularly if reviewers also are from prestigious universities, Benos et al. (2006) conclude that the empirical evidence in this respect is conflicting. Nevertheless, they note that the logic behind concealing the author’s identity is sound. So, it seems that the decision to adopt double-blind review rather than single-double review has taken place based on the logic rather than on empirical evidence. The question is about a precaution against a possible bias.

There is, of course, a challenge with the double-blind review to keep the author’s identity secret, if a reviewer really wants to know it. Assuming that the authors are properly concealed in the submission, my advice to reviewers is that do not make any effort to find out who the authors are. It makes fair and objective reviewing easier to you.

In the case of revealing reviewers’ identities to the authors, the evidence is also conflicting. On one hand, it has been argued that if reviewers’ identities are revealed to authors, they likely provide more objective, fair and developmental reviews, since they must be prepared to defend their reviews publicly (Hillman and Rynes 2007). Benos et al. (2006) on the other hand conclude that revealing reviewers’ identities has no beneficial effects in terms of substantive quality or strength of their reviews, but the authors may feel that they receive more courteous and constructive feedback. So, if there is a risk of bad reviews because of reviewer anonymity as Armstrong (1982) and Hillman and Rynes (2007), for example, suggest, why don’t we attempt to safeguard against it as a precaution?

Benos et al. (2006) point out that if reviewers’ identities are not concealed, reviewers more likely decline to review. They refer to an experimental study of Goodlee et al. (1998) in which 50 % of the reviewers declined to participate when asked to sign their reviews, while 46 % declined in groups in which no signing was requested. So, this does not indicate any significant decrease in the readiness to review when requested to sign. Yet, van Rooyen et al. (1999) report a 12 % decrease in the readiness to review when reviewers are requested to sign compared with the situation in which their identities are blinded to the authors.
Regehr and Bordage (2006) conducted a survey of reviewing preferences among 838 authors and reviewers of Medical Education journal. They found that about 50 % of reviewers resisted the idea that their names are revealed to the authors. The remaining 50 % either preferred that they will be revealed (roughly 22 %) or were indifferent (roughly 28 %). Those who resisted mentioned a number of reasons for their position: to facilitate honest reviewing, to avoid ‘bad blood’ and tensions among colleagues and friends, including interfering career development and grant applications, saving authors from knowing the identity of a reviewer returning harsh critique on one’s work.

The risk that a critical report – even when developmental – may negatively influence a reviewer’s future career development and grant applications may be a valid concern especially among junior reviewers (who do not have tenure, for example). Yet, I have problems to understand that a senior reviewer (with tenured professorship, for example) does not have courage to write honest criticism, even when his/her name is revealed to the authors whoever they are. Even when highly critical, a review can be written in a polite and developmental way.

Personally, I have never had any problem with the idea that my name would be revealed when I have reviewed. If I have attempted to be as honest, fair, objective and constructive as possible as a reviewer, why should I conceal my identity from the authors? It may be partly because of my cultural background, in which it is not much appreciated if a person tells one story in the front of you and a totally different story behind your back.

Those of you, who agree with me, please consider signing your reviews. Hillman and Rynes (2007) refer to McCook’s (2006, p. 3) quote of the editor of Journal of the American Medical Association as saying, “I’ve always signed every review I’ve ever done, because I know if I sign something, I’m more accountable. Juries are not anonymous, neither are people who write letters to the editor, so why are peer reviewers?” Unfortunately, I have never done it, since in hindsight it may be the biggest mistake during my academic career. Perhaps I would have more academic friends, if I had done so.

Actually I have an example in which concealing reviewers’ identities likely caused ‘bad blood’. The question is about a highly cited article with well above 1000 ISI Web of Knowledge™ citations. For some reason I have sensed that one of its authors has always been exceptionally reserved in my company, when we have occasionally met in different roles during the years. Later I got to know that this seminal article had enormous difficulties to get accepted. When I heard about that, it occurred to me that perhaps this specific author has assumed that I have served as a reviewer of his article. This article refers to a few of my publications. It may well be that one of its reviewers has suggested these articles in his/her review report. If so, this reviewer must have been quite knowledgeable of my work at that time. But (s)he was not me.

In conclusion, I am strongly in favor of single-blind reversed review in which the authors are concealed but reviewers are revealed, at least in its optional form. This optional form allows that junior reviewers can choose whether their names are revealed.
or not, but more senior reviewers’ names are revealed at least when the review process is closed.

Some IS journals such as *Journal of the AIS* and *MIS Quarterly* have applied a practice in which the names of reviewers are optionally revealed in the accepted paper, if the reviewers are ready for that. However, it seems that both journals have abandoned this practice. It may not be any big loss, since revealing reviewers of an accepted paper is not usually a big problem, if a reviewer has not been highly critical in its case or if the paper does not turn out to be a Type I error. Rejected papers are much more problematic, especially in the case of Type II errors, when it later turns out that an excellent paper has been rejected.

**Conclusions**

This essay suggested three system-level means to promote better reviewing and to complement individual-level guidelines such as Lee (1995). The suggested means are fairly independent of each other so that they can be adopted individually in any order or all together, even though I believe that together they are most effective in improving the quality of peer reviews.

In the case of first proposal, I would expect that most reviewers would welcome systematic feedback about their reviews. It does not require much additional work. If there is a short questionnaire for authors to collect that information after each review cycle, they do not have much choice but to answer. The challenge is to convince them that it is confidential during the ongoing review process so that it will not interfere with it and that they can decide how it is delivered (non-anonymously or anonymously, immediately after the process is over, or as a part of aggregated feedback package after the review has a sufficient number of reviews).

I would expect that rewarding active and good reviewers would also be well received at least by such reviewers. A critical question is how the editors-in-chief take this proposal. As argued above it does not necessarily lower the quality of journals when compared with issues and opinions, research essays, research notes and similar contributions. Furthermore, if it motivates reviewers to do better job, it gradually leads to accepted papers of improved quality.

Generally, I suppose that the suggestion to reveal reviewers’ identities is the most difficult to accept since the reviewer anonymity may be some sort of taboo. I understand that there is some resistance to the idea. Yet, it is encouraging that Regehr and Bordage (2004) found that about 50% of senior reviewers seem not to totally exclude it. Furthermore, I guess that part of the resistance is an outcome of the existing practice. For example, if it is regarded as a sign of seniority and potential for editorial positions that a reviewer has courage to reveal their identities to the authors, the situation may change.

These suggestions are based on my own experiences as an author, reviewer and editor and especially by two cases of quite bad reviews I have encountered. Yet, the two cases ended fairly happily from my viewpoint. Soon after the rejection of my ISJ submission I was invited to join the editorial board of ISJ. I have never asked from David Avison and Guy Fitzgerald, the editors-in-chief of ISJ at that time, whether this
unfortunate review was the reason for the invitation. I suppose so. In any case it was a great honor, my first editorial board position in a premier IS journal, and perhaps opened the door for additional ones. So, if you receive a really bad review, do not get depressed. Perhaps, as an apology you will be invited to join the editorial board of the journal.

In the second case, I was provided an opportunity to revise. I did it just in order to provide in my response letter straightforward feedback to the reviewer in question. I did not expect anything from the second review after my response. I do not know if the reviewers remained the same, but I suppose that some of them were new. The quality of these reviews was acceptable, even though my manuscript was rejected. Yet, I was provided an opportunity to submit a shortened version as an Issues and Opinions paper. I did so and the manuscript was eventually accepted (Iivari 2014). I suppose that it was a sort of reward for my previous services to EJIS, a kind of “reviewer forum” article even though no categorized in that way. That is fine for me.

During the former case I was in the middle of my career and acceptance or rejection of one paper did not make any big difference to me. During the latter case I had just retired and the rejection had still less impact on my career, although one always hopes that one’s manages to publish on a good forum so that the ideas receive wider publicity. At that stage I did not see any reason to silently tolerate a bad review. I made my perception clear in my response letter.

A sad conclusion of the latter story is that I sensed the editors involved were quite defensive when I criticized the quality of one particular review. Rather than admitting the situation, they simply denied it. The official view throughout the review process was that it was a review of normal quality. My criticism of the quality of the review was against the implicit rules of conduct. Authors are not supposed to do so.

Generalizing from my experience, one cannot trust that the editors are ready to eliminate really bad reviews. Therefore journals should have a system that firstly encourages to write good and developmental reviews, makes it possible to identify those who repeatedly continue to write weak ones, and directs journals into corrective actions when a reviewer continuously keeps on writing such reviews. I wish that the three system-level suggestions outlined above are able to improve the situation.

References

Armstrong, J., Research on scientific journals: implications for editors and authors, *Journal of Forecasting*, 1, 1982, pp. 83–104
Brown T., *Peer Review and the Acceptance of New Scientific Ideas*, Sense About Sci-


Hevner, A.R., March, S.T., Park, J. and Ram, S., Design science in information systems research, MIS Quarterly, 28(1), 2004, pp. 75-105


March, S.T. and Smith, G.F., 1995, Design and natural science research on information technology, Decision Support Systems 15, pp. 251-266


Regehr G. and Bordage, G., To blind or not to blind? What authors and reviewers prefer, Medical Education, 40, 2006, pp. 832-839

Saunders C., From the trenches: thoughts on developmental reviewing, *MIS Quarterly*, 29(2), 2005b, pp. iii-xii
Santana, A.D., Virtuous or vitriolic: The effect of anonymity on civility in online newspaper reader comment boards, *Journalism Practice*, 8(1), 2014, pp. 18-33
Straub, D., Editor’s comments: Type II reviewing errors and the search for exciting papers, *MIS Quarterly*, 32(2), 2008, pp. v-x