

1-2016

Practical Suggestions for Improving Scholarly Peer Review Quality and Reducing Cycle Times

Paul Ralph

University of Auckland, paul@paulralph.name

Follow this and additional works at: <http://aisel.aisnet.org/cais>

Recommended Citation

Ralph, Paul (2016) "Practical Suggestions for Improving Scholarly Peer Review Quality and Reducing Cycle Times," *Communications of the Association for Information Systems*: Vol. 38 , Article 13.

DOI: 10.17705/1CAIS.03813

Available at: <http://aisel.aisnet.org/cais/vol38/iss1/13>

This material is brought to you by the Journals at AIS Electronic Library (AISeL). It has been accepted for inclusion in Communications of the Association for Information Systems by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.



Practical Suggestions for Improving Scholarly Peer Review Quality and Reducing Cycle Times

Paul Ralph

Department of Computer Science
University of Auckland, New Zealand
paul@paulralph.name

Abstract:

Scholarly peer review is both central to scientific progress and deeply flawed. Peer review is prejudiced, capricious, inefficient, ineffective, and generally unscientific. Management journals have longer review cycles than journals in other fields. Long cycle times demonstrably harm early-career researchers. Meanwhile, a lack of transparency conceals and facilitates editorial misconduct, and some dismiss legitimate criticism of peer review as unfounded resentment. We can address these problems by eliminating unnecessary reviewing, simplifying the peer review process, introducing author rebuttals, creating an AIS ombudsman, and enforcing the relationship between submitting and reviewing. These problems are, however, entangled with fundamental problems with journals. Ultimately, therefore, we can only fix peer review in conjunction with replacing journals with repositories.

Keywords: Peer Review.

This manuscript was solicited by the Department Editor for Debates, Karlheinz Kautz.

1 Problems with Peer Review

Scholarly peer review is "the lynchpin about which the whole business of science is pivoted" (Ziman, 1968, p. 111). It determines what grants are awarded, where papers are published, and which scientists are tenured and promoted (Smith, 2006). By making some topics and methods easier to publish, peer review indirectly determines what is studied and how.

Yet, peer review is deeply flawed in at least the following six ways.

1. Peer review is biased. Peer review is demonstrably sexist (Figure 1) and nepotistic (Wenneras & Wold, 2001). It is biased against novel ideas (Campanario, 1995). As Mahoney (1997, p. 161) state: "Reviewers [are] strongly biased against manuscripts which [report] results contrary to their theoretical perspective". Many times, including just this week, I have been told that I would need to collaborate with some senior people to get a proposal or paper accepted. The reader has probably given or received the same advice. If success is based on who is involved rather than the quality of the work, it is crucial to call this what it is: corruption.

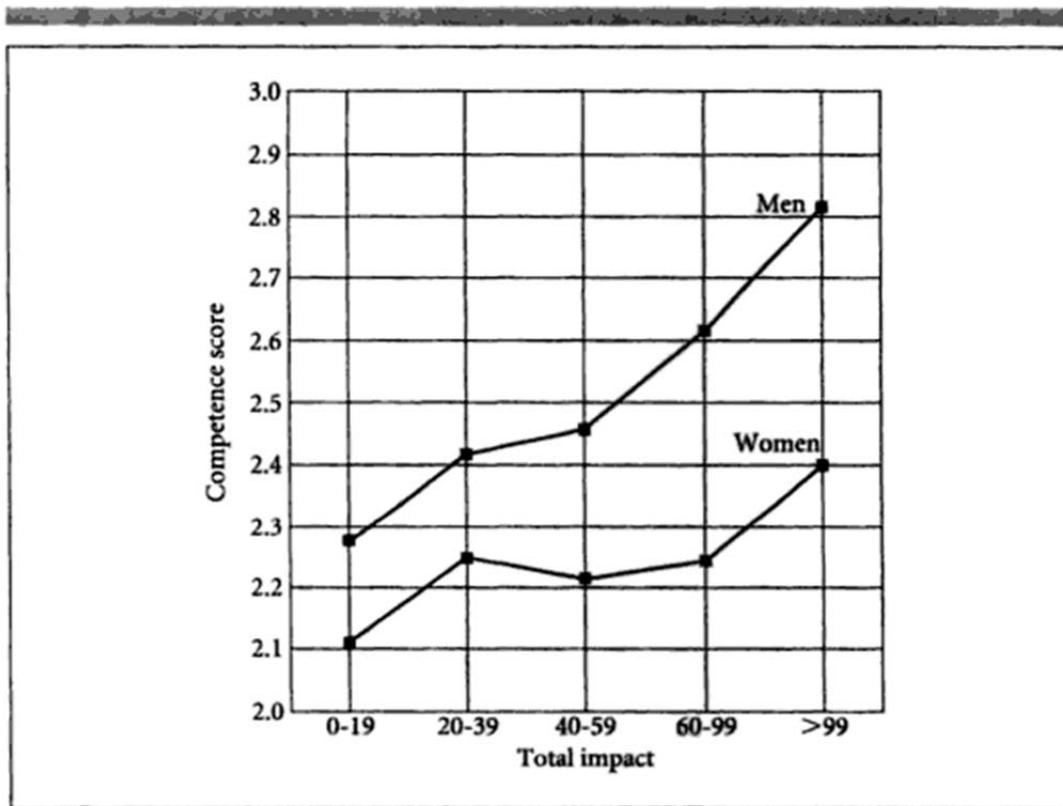


Figure 1. Mean Competence Scores Assigned by Grant Application Reviewers to Men and Women with Equivalent Publication Records (Adapted from Wenneras & Wold, 2001, p. 48)

2. Peer review is capricious and ineffective. Inter-rater reliability is so low the peer-review is mostly random (Cole, Cole & Simon, 1981; Peters & Ceci, 1982; Lock, 1991; Rothwell & Martyn, 2000). Peer review, therefore, does not ensure quality. Peer review also cannot reliably detect fraud (Brown, 2004).
3. Peer review lacks construct validity. As Bohannon (2013, p. 64) state: "Most reviews [focus] exclusively on the paper's layout, formatting, and language". Peer review is intended to evaluate a study's methodological correctness (Campanario, 1995); however, the reviews I have received from top MIS journals and conferences predominately focus on framing, positioning, and tone—they largely ignore methodology. Worse, many reviews are emotionally destructive. My first peer review (when I was 19 years old) encouraged me to "consider a different career".
4. Peer review is unreasonably slow and management journals are significantly slower than other fields (Figure 2). The mean time from submission to publication across fields is about one year (Björk & Solomon, 2013). While we might expect slower review cycles in management than

physics or chemistry because management has less methodological consensus, this does not explain why management is slower than social sciences or arts and humanities.

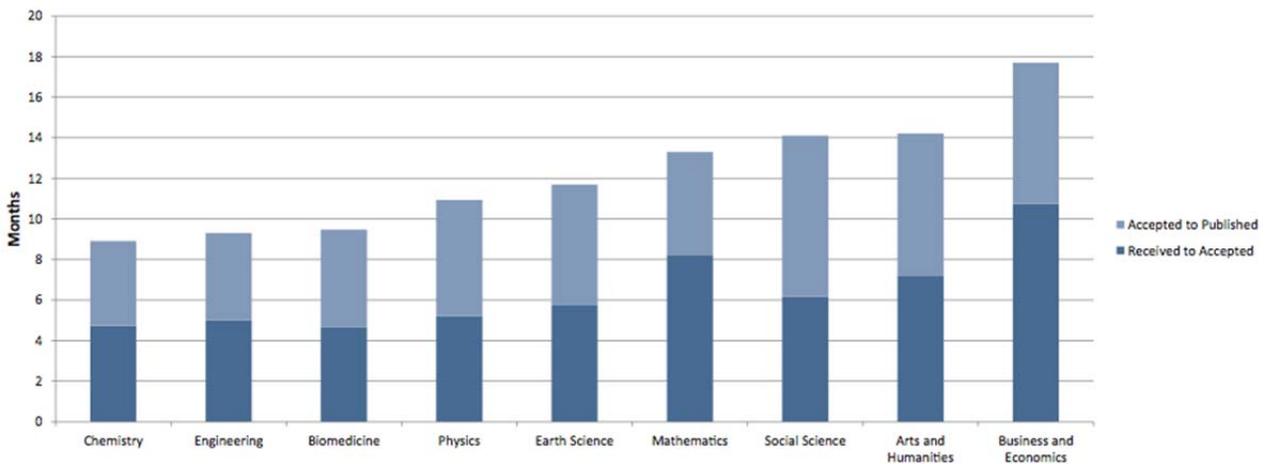


Figure 2. Average Publication Times in Months by Field (from Bjork & Solomon 2013)

5. Peer review conceals editorial misconduct. As Gollogly and Momen (2006, p. 24) state: “Editorial misconduct includes failure to observe due process, undue delay in reaching decisions and communicating these decisions to authors, inappropriate review procedures, and confounding a journal’s content with its advertising or promotional potential... failure to investigate suspected misconduct, failure to retract when indicated, and failure to abide voluntarily by ... international guidelines”. While little research has investigated editorial misconduct, in my short career, I have witnessed an astounding number of cases. The most dramatic involved a paper that criticized a standard being assigned to and desk rejected by an associate editor (AE) who co-authored the standard.
6. Some dismiss legitimate criticism of peer review as sour grapes.

livari (2016) makes three suggestions for improving peer review quality: give reviewers systematic feedback, reward good reviewers with publication opportunities, and reveal reviewers’ identities. While I agree with unmasking reviewers, livari’s suggestions neither acknowledge nor address the six core problems that I outline above. Providing systematic feedback to reviewers is more work and will, therefore, increase cycle times. Giving “good” reviewers special opportunities to publish in a system known for nepotism, sexism, and bias would be rife for abuse. Unmasking reviewers does not improve review quality or acceptance rates (McNutt, Evans, Fletcher, & Fletcher, 1990; Justice et al., 1998); however, masking authors may improve review quality (McNutt et al., 1990; Jadad et al., 1996). Worse, livari downplays long review cycle times as “thorough reviewing takes time”. There is no relationship between review quality and time taken to review (van Rooyen, Godlee, Evans, Black, & Smith, 1999), and most reviews take less than four hours (Evans, McNutt, Fletcher, & Fletcher, 1993). Furthermore, an emeritus professor’s dismissing publication delays must be especially galling to academics denied tenure over insufficient publications.

We de-bias judgments by redesigning the person-task system to mitigate the bias (Fischhoff, 1982; Larrick, 2008), not by giving slow, inconsistent feedback. We fix an instrument exhibiting poor construct validity and inter-rater reliability by changing the questions, not by incentivizing the judges. We accelerate a process by making it simpler, not more complicated.

2 Recommendations for Improving Peer Review

Peer review has evolved substantially over time (Weller, 2001) and can yet be improved and simplified by transitioning to a process like that in Figure 3. In this section, I describe the key aspects of this process and how it addresses review quality and cycle times.

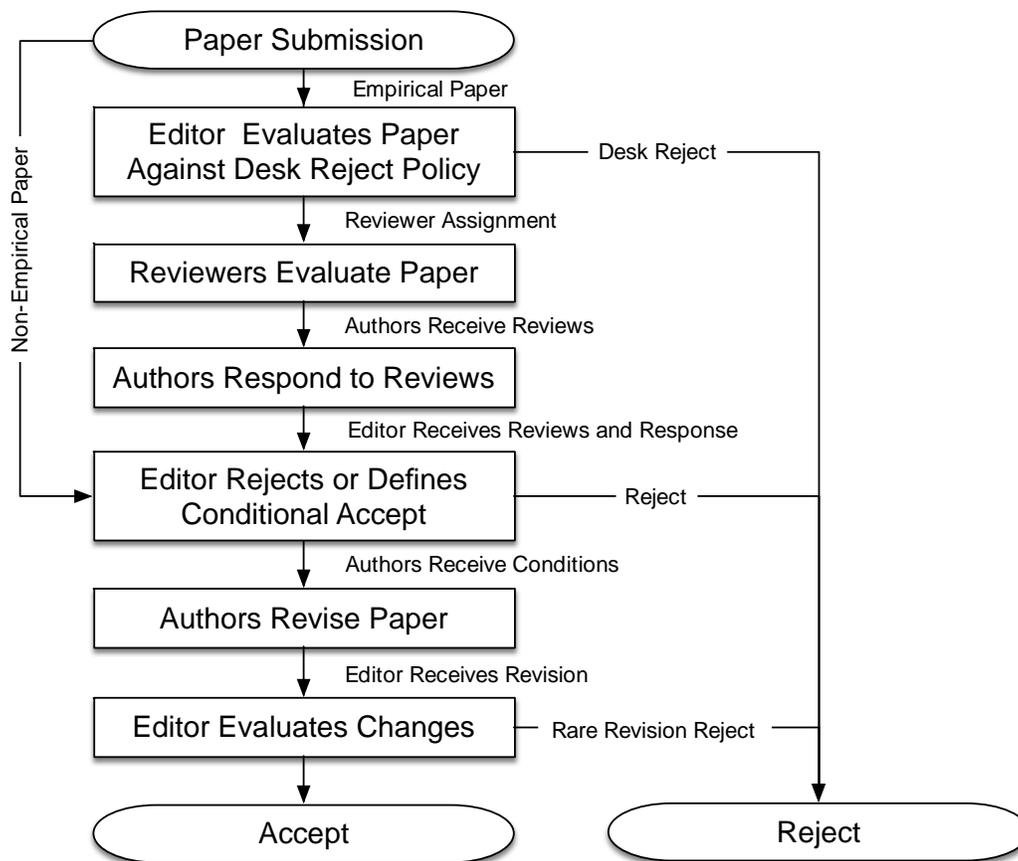


Figure 3. A More Effective and Efficient Peer Review Process

2.1 Reduce the Number of Reviews

Finding reviewers is difficult because journals unnecessarily review too many papers. Position papers, opinion pieces, editorials, notes, workshop reports, teaching cases, experience reports, conceptual explorations, non-systematic literature reviews, and papers that provide guidelines for conducting, evaluating, and presenting research do not require peer review. Peer review is scientific/methodological critique, not literary criticism. If a paper does not have a methodology, it does not need peer review. Journals' editors may simply evaluate these papers.

Another way to reduce the reviewing burden is to develop a clear set of desk-reject rules that allow more desk rejects. For example, we might desk reject any interview-only study based on less than 10 hours of interviews, any experimental study that does not clearly indicate its independent and dependent variables, or any paper that does not discuss the assumptions of its statistical tests. If such policies are published, challenged, and continually evolved, they save everyone time and drive up research standards.

Furthermore, papers do not need multiple editors. One can reduce the reviewing burden by flattening the editorial hierarchy and assigning each paper to a single editor. Having both senior editors and associate editors unnecessarily complicates and lengthens the review process for no demonstrable benefit. Younger scholars are better at evaluating manuscripts (Evans et al., 1993); consequently, assigning editorial roles based on seniority is irrational.

Moreover, papers do not need multiple rounds of peer review. If one asks reviewers direct questions about the methodological substance of a paper rather than its style, one round of review is sufficient. Editors are entirely capable of deciding whether a revision adequately addresses reviewers' comments. Limiting peer review to one round will make it easier to recruit reviewers because they are agreeing to less work, will improve review quality because reviewers will know they only have once chance to give a comprehensive evaluation, will reduce late-stage rejects (which are devastating to early-career academics), and will discourage reviewers from micromanaging papers.

This approach contradicts the recent trend toward “developmental” peer review. Developmental review is paternalistic and counterproductive. When someone submits a paper for review, the research is done. There is nothing left to develop except framing, positioning, and tone, which are neither central to science nor the purview of peer review. While a paper may need to tone down some claims or discuss a few limitations, spending three rounds of review arguing about wording is an unconscionable waste of time and resources. Developmental review increases cycle times, allows reviewers to micromanage papers, and leads to devastating late-round rejections. There is simply no evidence that any tangible benefit outweighs these costs. Treating peer review as cross-sectional evaluation rather than longitudinal development will reduce cycle times and encourage reviewers to focus on methodological issues.

2.2 Review the Study, not the Paper

Peer review exhibits low quality and reliability because editors ask the wrong questions. If a paper uses structural equation modeling to evaluate data from a questionnaire about e-governance, the editor should specifically ask:

- An expert in structural equation modeling to evaluate the data analysis
- An expert in questionnaire methodology to evaluate instrumentation and data collection, and
- An expert in e-governance to evaluate contribution and construct validity.

Editors should contrastingly not ask reviewers to recommend a decision, evaluate presentation issues, or complain about missing literature.

When one evaluates a study rather than a paper, only two decisions are evident:

1. Reject because the study has unacceptable limitations or flaws, or
2. Accept, possibly conditional on updated data analysis or dialing back some of the paper’s claims.

There is simply no room for a major revision or, even more absurd, a “reject but invite resubmission”. If more data collection is not necessary, revisions are minor. If more data collection is necessary, then it is beyond the scope of the review process and warrants a reject.

This approach has many benefits. It will accelerate peer review because it is much simpler and faster to evaluate just one aspect of a paper. It will make it easier to find reviewers not only because reviewing will take less work but also because reviewers are less likely to perceive a paper as outside their expertise. It will improve review reliability because questions such as “is this sampling strategy reasonable?” are more specific and less controversial than questions such as “should this be published in this journal?”. By not asking reviewers for recommendations, editors will have to consider the reviewers’ analyses rather than simply passing on the majority decision. It will improve review quality by focusing reviewers on the substance over style. No reviewers should be asked about positioning, framing, tone, the phrasing of research questions, or whether they are offended that the author has not referenced enough of their papers. No section should be provided for such comments and, if they are worked into other sections, they should be ignored. While it is unfortunate to ignore reviewers’ possibly helpful comments, it may be the only way to get reviewers to focus on substance.

Furthermore, since editors are peers, if the editor is an expert in one or more of these areas, inviting a reviewer in that area may be unnecessary. Editors should invite reviewers to fill in specific gaps in their knowledge and not simply for another opinion.

2.3 Invite Author Rebuttals and Appeals

Another way to improve the peer review process to give authors an opportunity to respond to reviewer comments before the editor makes a decision. Reviewers are not perfect. In many cases, negative reviews stem from easily correctable misunderstandings. I once had a paper rejected primarily because it lacked detail of how the proposed system worked. In fact, the paper’s appendix provided the complete algorithm, which the reviewer missed. One can easily rectify this kind of honest mistake by allowing authors to rebut reviews. Moreover, pre-decision responses put the editor in a more judge-like role: they weigh reviewers’ prosecution against authors’ defense. Rebuttals are increasingly popular with top conferences including CHI. While this adds a step, it can save time, limit frustration, reduce late rejects, and increase fairness and transparency.

To address editorial misconduct, authors need a way to formally appeal editorial decisions outside the existing editorial chain of command. I have reported numerous counts of editorial misconduct. In my experience, no conflict of interest, process malfunction, editorial error, or reviewer comment is sufficiently egregious to elicit intervention from an editor-in-chief. Even the AE conflict of interest I mention above was casually dismissed. The Association for Information Systems should, therefore, create an ombudsman's office to investigate editorial misconduct. The office should clearly define and publicize grounds on which one can make a complaint, which includes conflict of interest, unreasonable delay, failure to retract, capricious use of evaluation criteria, factual errors, ad hominem attacks and outright nonsense.

2.4 Timeboxing and Quotas

When authors submit an article to an outlet, they implicitly agree to review some papers for that outlet independently of other editorial or service roles. We can mostly solve the problem of finding reviewers can by simply explicating and enforcing this agreement. If someone refuses a review, fails to complete it on time, or does a terrible job, the journal should put their in-press papers on hold and suspend them from submitting papers for a year. The journal should publish their names and publicly shame them. This may come with the added benefit of reducing gift authorship because certain senior professors will be inundated with reviewing. Of course, reviewing should be limited to two or three reviews per submission and suspensions should be transparent—silently blacklisting people contributes to nepotism. Moreover, authors still should be able to refuse a review over a conflict of interest or medical reason.

With this quota system in place, outlets can timebox each step in the review process. In project management, timeboxing not only prevents analysis paralysis but also trains participants to finish tasks on time (Schwaber, 2004). Timeboxing a reviewing period to, for example, one month, means that the editor must make a decision with whatever reviews are completed by the end of the month. Timeboxing the author rebuttal period similarly means that, on the deadline, the editor has to make a decision whether or not the authors have provided a rebuttal. Timeboxing represents a commitment to finishing tasks on time and is central to reducing cycle times.

3 The End of Journals

More fundamentally, the journal is a 19th century solution to a 21st century problem. We no longer need curated content printed on dead tree and delivered by ship. Journals (and their rankings) intrinsically undermine scientific progress in two ways:

1. Journals exacerbate the file drawer effect: that is, selective publication biases the scientific literature toward dramatic, positive findings. The file drawer effect hinders meta-analyses and, therefore, evidence-based decision making, which has led to numerous calls for researchers to simply publish all empirical trials (Chalmers, Glasziou, & Godlee, 2013).
2. Journal rankings create an elaborate, repeated competition that inefficiently and capriciously determines which papers are published in which journals. Both authors' and reviewers' efforts are wasted every time a paper is rejected from one reputable journal only to be published in another. One could argue that each review cycle improves the paper but this is simply untrue (Peters & Ceci, 1982), not least because peer review is practically random and, by the time the paper is submitted, it is too late to correct most serious problems (see above). This is an unconscionable waste of time and resources for no benefit. Journal rankings moreover facilitate the neoliberal assault on academia (Barkawi, 2013) by creating yet another oversimplified, misleading, invalid metric by which to measure and manipulate academics. The quality and impact of research is not determined or measured by where it is published. Fortunately, the impact of publication venue on future citation is decreasing (Acharya et al., 2014).

Yet, journals do attempt to:

- Archive scholarly work for the long term
- Promote research on particular subjects to academics interested in those subjects
- Copyedit and typeset papers, and
- Indicate the legitimacy of scholarly work to non-scientists including journalists.

Consequently, “abandoning journals altogether, in favor of a library-based scholarly communication system, will ultimately be necessary” (Brembs, Button, & Munafò, 2013, p. 1), which means creating

repositories such as arXiv.org that are responsible for permanently archiving any empirical scholarly work. One can organize repositories by subject such that academics can subscribe to updates from relevant subjects. They can employ copyeditors and typesetters. Rather than accepting or rejecting manuscripts, repositories can use peer review to summarize studies and their limitations aimed at non-scientists. It is also possible—although not necessarily wise—to have the review process produce grades for each paper. For example, reviewers could rate a paper on two separate scales, methodological quality and contribution/originality. Such a system would have many benefits, not least emancipating scholarly publication from commercial publishers.

4 Conclusion

In summary, scholarly peer review is not somewhat imperfect but catastrophically broken. Fixing it will require fundamental changes, not fiddling at the margins. Empirical research has consistently shown that peer review is prejudiced, capricious, inefficient, ineffective, and unscientific. Management, including MIS, has longer review cycles than other fields, which demonstrably harms early-career researchers. Meanwhile, lack of transparency conceals and facilitates editorial misconduct, and, worst of all, some dismiss legitimate criticism of peer review using ad hominem attacks. We can address these problems in the short term by eliminating unnecessary reviewing, simplifying the peer review process, introducing author rebuttals, creating an AIS ombudsman, and enforcing the relationship between submitting and reviewing. In the long term, we can only overcome the problems with peer review by replacing journals with repositories.

References

- Acharya, A., Verstak, A., Suzuki, H., Henderson, S., Iakhiaev, M., Lin, C. C. Y., & Shetty, N. (2014). Rise of the rest: The growing impact of non-elite journals. *arXiv: 1410.2217*.
- Barkawi, T. (2013). The neoliberal assault on academia. *Al Jazeera*. Retrieved from <http://www.aljazeera.com/indepth/opinion/2013/04/20134238284530760.html>
- Björk, B.-C., & Solomon, D. (2013). The publishing delay in scholarly peer-reviewed journals. *Journal of Informetrics*, 7(4), 914-923.
- Bohannon, J. (2013). Who's afraid of peer review? *Science*, 342(6154), 60-65.
- Brembs, B., Button, K., & Munafò, M. (2013). Deep impact: Unintended consequences of journal rank. *Frontiers in Human Neuroscience*, 7.
- Brown, T. (2004). *Peer review and the acceptance of new scientific ideas*. London, UK: Sense about Science.
- Campanario, J. M. (1995). Commentary on influential books and journal articles initially rejected because of negative referees' evaluations. *Science Communication*, 16, 304-325.
- Chalmers, I., Glasziou, P., & Godlee, F. (2013). All trials must be registered and the results published. *BMJ*, 346, f105-f105.
- Cole, S., Cole, J. R., & Simon, G. A. (1981). Chance and consensus in peer review. *Science*, 214(4523), 881-886.
- Evans, D. A. T., McNutt, R. A., Fletcher, S. W., & Fletcher, R. H. (1993). The characteristics of peer reviewers who produce good-quality reviews. *Journal of General Internal Medicine*, 8(8), 422-428.
- Fischoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge, MA: Cambridge University Press.
- Gollogly, L., & Momen, H. (2006). Ethical dilemmas in scientific publication: Pitfalls and solutions for editors. *Revista de Saúde Pública*, 40, 24-29.
- Iivari, J. (2016). How to improve the quality of peer reviews? Three suggestions for system-level changes. *Communications of the Association for Information Systems*, 38, 263-273.
- Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J. M., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials*, 17(1), 1-12.
- Justice, A. C., Cho, M. K., Winker, M. A., Berlin, J. A., Rennie, D., & Investigators, T. P. (1998). Does masking author identity improve peer review quality? A randomized controlled trial. *JAMA*, 280(3), 240-242.
- Larrick, R. P. (2008). Debiasing. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 316-338). Malden, MA: Blackwell.
- Lock, S. (1991). *A difficult balance*. Oxford: Wiley-Blackwell.
- Mahoney, M. J. (1977). Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research*, 1(2), 161-175.
- McNutt, R. A., Evans, A. T., Fletcher, R. H., & Fletcher, S. W. (1990). The effects of blinding on the quality of peer review: A randomized trial. *JAMA*, 263(10), 1371-1376.
- Peters, D. P., & Ceci, S. J. (1982). Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behavioral and Brain Sciences*, 5, 187-195.
- Rothwell, P. M., & Martyn, C. N. (2000). Reproducibility of peer review in clinical neuroscience. *Brain*, 123, 1964-1969.
- Schwaber, K. (2004). *Agile project management with Scrum*. Redmond, WA: Microsoft Press.
- Smith, R. (2006). Peer review: A flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine*, 99(4), 178-182.

- van Rooyen, S., Godlee, F., Evans, S., Black, N., & Smith, R. (1999). Effect of open peer review on quality of reviews and on reviewers' recommendations: A randomised trial. *BMJ*, 318(7175), 23-27.
- Weller, A. C. (2001). Editorial peer review: Its strengths and weaknesses. Medford, NJ: *Information Today*.
- Wenneras, C., & Wold, A. (2001). Nepotism and sexism in peer-review. *Nature*, 387(6631), 341-343.
- Ziman, J. M. (1968). *Public knowledge: An essay concerning the social dimension of science*. Cambridge, UK: Cambridge University Press.

About the Authors

Paul Ralph is an author, scientist, consultant, and computer science lecturer at The University of Auckland. His research centers on the theoretical and empirical study of software and game development, including projects, processes, practices, tools and developer cognition, socialization, productivity, creativity, wellbeing and effectiveness. His research has appeared in premier software engineering and information systems outlets, including the International Conference on Software Engineering and the *Journal of the Association for Information Systems*. Additionally, he has written editorials on technology, education and design for influential outlets including Business Insider, Lifehacker and The Conversation. He co-founded the AIS Special Interest Group for Game Design and Research (SIGGAME). He holds a PhD in Information Systems from the University of British Columbia. Previously, he was a lecturer at the Lancaster University Management School, the highest rated management research institution in the United Kingdom.

Copyright © 2016 by the Association for Information Systems. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and full citation on the first page. Copyright for components of this work owned by others than the Association for Information Systems must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or fee. Request permission to publish from: AIS Administrative Office, P.O. Box 2712 Atlanta, GA, 30301-2712 Attn: Reprints or via e-mail from publications@aisnet.org.