More on Improving Reviewing Quality with Double-Blind Reviewing, External Review Committees, Author Response, and in Person Program Committee Meetings

Kathryn S McKinley, Principal Researcher, Microsoft Research

June 2015

Peer review is the central mechanism that scientists use to evaluate research manuscripts, artifacts, and individual researchers, and thus advance the state of scientific knowledge and determine career trajectories. Peer reviewing seeks to evaluate research using *qualified experts*, who have the scientific training, experience, and skills to judge the merit of the work. Just as scientific progress and methods are evolving, so is the peer review process. Conferences are the primary publication venues for computer scientists. Current challenges for peer reviewing in computer science include (1) obtaining expert reviews as the breadth and depth of our field grows; (2) accommodating growth in the amount of research, number of researchers, and submissions; and (3) incorporating best practices based on social science research on nepotism and bias.

In good news, the SIGPLAN and ACM communities have substantially improved peer reviewing practices of conference submissions over the past decade. In 2008, I wrote a SIGPLAN Notices Editorial on reviewing practices that advocated for external review committees, author response, and double-blind reviewing, none of which were common in SIGPLAN conferences at that time [14]. *External review committees* consist of reviewers, who have the same accountability and responsibilities as the program committee, but generally review fewer papers and do not travel to committee meetings. They are intended to deliver high expertise reviews and to accomodate increasing numbers of submissions. *Author response* gives authors a single opportunity to answer reviewer questions and correct misunderstandings before the committee decides the fate of the submission. It is intended to increase reviewer accountability. With *double-blind reviewing,* neither authors nor reviewers are revealed. With single-blind reviewing, authors are revealed, but reviewers are not. Double-blind reviewing is intended to reduce implicit reviewer bias with respect to individual researchers, research institutions, race, and gender. These policies are now widely used.

In this article, I review the motivation and implementation issues for reviewing best practices with a focus on conferences. I revise and expand on some of my prior recommendations [14], reflecting changes in the community and my own experiences as program chair (ASPLOS, PLDI, PACT, ISMM, CGO), Editor of TOPLAS, and 50+ committee reviewing experiences. In particular, I advocate for (1) double-blind until accept to further reduce bias and enable rejected submission to remain blind in subsequent submissions; (2) the need for external review committees (vs ad hoc assignments) to handle increasing numbers and breadth of submissions, ease the implementation of double-blind reviewing, and improve overall review quality with more expert reviews; (3) the use of author response to improve reviewer accountability; and (4) in person program committee meetings to communicate, establish, and constantly apply scientific standards and build community. While (2) and (3) are now common, (1) double-blind implementation remains varied, and I believe (4) is undervalued as a mechanism for improving community and reviewing accountability (trust and verify).

1. Introduction

The most important job of a program chair, program committee, and external review committee is to deliver expert, fair, and thoughtful reviews. Because scientific progress depends on peer reviewing, it behooves researchers to ensure that peer review evaluations come from qualified experts, equipped to evaluate contribution, novelty, correctness, and methods, and that reviewers have sufficient time to provide evaluations. Because of the breadth of the field, it may take more than one reviewer to deliver expertise for the entire submission. Delivering sufficient reviewer expertise for every submission is a responsibility of all reviewers, and the primary job of the program chair.

Researchers are human with emotions and biases. The reviewing process acknowledges our humanity. In particular, *single-blind* reviewing never reveals the reviewers' identity to the authors, in order to encourage fair and frank reviewing, and protect reviewers who recommend rejecting submissions from author retribution. All the conference, journal, and grant processes of which I am aware use at least single-blind reviewing. Many now use *double-blind* reviewing. In addition to not revealing reviewer identities, the authors' identities are not known to the reviewers for all or most of the reviewing process in double-blind. Its purpose is to focus the evaluation on the quality of the submission by reducing human biases with respect to the authors' race, gender, institution, and reputation, by not revealing those details.

Compared to single-blind reviewing, many studies show that double-blind reviewing improves the result quality and fairness. Many ACM conferences sponsored by SIGPLAN, SIGARCH, SIGMETRICS, SIGMICRO, and SIGMOD [1, 6], some computer science journals, such as TODS [4], and many journals in other disciplines, such as *The Journal of Finance* [2], successfully use double-blind reviewing.

The SIGPLAN PLDI and POPL communities significantly prefer double-blind reviewing and prefer a double-blind until accept policy. In the 2007 PLDI attendee survey, 148 attendees responded from a total of 334, a 44% response rate. Respondents indicated double-blind reviewing was: very useful (10%), useful (50%), neutral (30%), not useful (14%), or harmful (4%). Only 19% were opposed and 60% support double-blind. A 2012 survey with 275 responses from the 494 authors of POPL 2012 submissions shows that 70% of authors prefer double-blind over single-blind [9]. In the most recent survey of authors from PLDI 2015 submissions, 95 responded. Of these, 58% preferred submission to remain double-blind until accept [10]. 40% strongly agreed, 18% agreed, 24% were neutral, 7% disagreed, and 10% strongly disagreed with the statement "I prefer a blind-until-accept policy." Both the research on bias and the community agree that double-blind is best.

I recommend that SIGPLAN require all its conferences and journals to use double-blind reviewing for evaluating research submissions, and furthermore that SIGPLAN advocates for an ACM wide policy that requires double-blind reviewing.

Below, I summarize and point to some of the literature and scientific studies on reviewing, discuss the types of biases that double-blind reviewing helps minimize, and suggest implementation strategies. These strategies include (1) author response, (2) an external review committee (instead of ad hoc external reviews and in addition to the program committee), and (3) double-blind until accept to limit bias and to maintain double-blind anonymity for rejected papers in future submissions.

2. Reducing Bias Improves Quality

Some *explicit* bias is inherent in the process. Researchers make scientific judgements based on their specific expertise and experiences, which imbibes them with explicit and implicit biases and makes them qualified to judge related research. Obtaining three to five reviews, exposing all reviews to all the reviewers, author response, and discussing reviews in public on line or in person, all serve to trust and verify that reviewers make appropriate scientific judgements based on their specific expertise.

A number of scientific studies have examined implicit bias due to nepotism and gender bias in the scientific evaluation processes [2, 8, 7]. I summarize three studies here and recommend Snodgrass 2006 for a more extensive literature analysis [3]. As a community of scholars and scientists, we all benefit if our conferences and journals publish rigorous high quality articles, evaluated using community best practices on originality, quality, and methodology. However, if our evaluations are influenced by implicit or explicit bias based on nepotism, gender, researcher reputation, or institution reputation, the quality of our science is degraded.

Unfortunately, men and women still express systematic bias against women. Consider the 2005 European Young Investigator Awards (EURYI) [7] and Wenneras and Wold's analysis of the 1995 Swedish Medical Research Postdoctoral Fellowship competition for biomedical research [8]. In the first EURYI competition, the European Science Foundation (ESF) awarded 3 of 25 (12%) fellowships to women, although 25% of applicants were women [7]. ESF has not provided further data for analysis, so it is possible the men were better. However, the data for the 1995 Swedish Medical Research (SMR) postdoctoral fellowship is

available. Wenneras and Wold forced SMR to provide the data by appealing to the Administrative Court of Appeal in Sweden, which ruled that the information fell under the Freedom of the Press Act [8]. In 1997, there were 114 applicants (62 men and 52 women) for 20 fellowships, which were awarded to 16 men and 4 women. They analyze applicant success rates based on gender, publication record, position in author list, quality of publication venue, quality of PhD granting institution, research area, and affiliations with selection committee members.

Wenneras and Wold found nepotism and gender bias were significant factors in the evaluation process. To be judged as good as their male counterparts, female applicants had to be 2.5 times more productive. For example, if you were a woman, you needed 3 more *Nature* or *Science* articles or 20 more articles in specialized, prestigious journals to be judged equal to a man. Although the SMR prohibited reviewers from evaluating applicants with which they had a conflict, e.g., their own PhD students or students from their institutions, that was insufficient to protect against nepotism. The other committee members systematically scored the applicants with relationships to other committee members higher. For example, if you had a conflict of interest with a committee member, you accrued an advantage of the equivalent of 3 *Nature* or *Science* articles compared to your peers. If the SMR committee had awarded fellowships without these biases, the quality of fellowship recipients would have improved dramatically.

A widely used, although imperfect, metric for research quality and influence is citation count, i.e., the number of publications that refer to a given article. Relative to other papers within a discipline, more highly cited papers are considered more influential. Laband and Piette use this quality metric to study single and double-blind reviewing outcomes [2]. They examine 28 economic journals which used both single (nonblinded) and double-blind (blinded) reviewing and find:

> Articles published in journals using blinded peer review were cited significantly more than articles published in journals using non-blinded peer review, controlling for a variety of author, article, and journal attributes.

They conclude that reviewers are better at applying objective criteria on submission quality with double-blind because the articles they accepted under this system have a higher citation rate than articles accepted using single-blind reviewing. There is every reason to believe these and other results on how human nature, gender biases, and racial biases affect the outcomes of scientific evaluation hold regardless of the venue and scientific discipline.

3. Advice for Authors on Blinding Submissions

Common sense and careful writing can easily preserve anonymity without detracting from the submission or dissemination of ideas.

Double-blind is not intended to retard dissemination of research. However, authors should never try to influence program committee members while their paper is under submission by advocating for their work or revealing their authorship. This ethical standard should hold regardless; double-blind simply adds revealing authorship as a form of influence. Authors should be permitted to post submissions and give talks on submissions. However, authors should not email, use social media, or otherwise explicitly act to bring their submissions to the attention of program committee members.

To make your submission double-blind, do not reveal the identity of any author in the text. For example, do not include author names, funding sources, or personal acknowledgments. Do not put your name in the submission document name; do not submit a file called McKinley.pdf whether or not your name is McKinley.

Never eliminate essential self-references or other references. Always use the third person when referring to your prior work. For example, if you are Smith write: "We build on the prior work by Jones and Smith [JS 2003]."

If you have a concurrent related submission, an institutional technical report, software, or other supplementary material that reveal your identity but that you believe reviewers will want to consider, send

it or the links to the Program Chair and/or add these supplementary materials as provided by the submission software. Reference this material as available from the program chair upon request.

If reviewers believe that double-blind is an obstacle to objectively reviewing a submission, they should contact the program chair, who has the option to unblind the submission. For example, If reviewers believe unblinded materials are essential to evaluating the research, the program chair may provide this information. This situation should be rare. Even following these guidelines, closely building on your own prior work may reveal author identity. However, without author names on submissions, reviewers are reminded their judgements should be restricted to the submission, not the authors. **Double-blind is not perfect, just better.**

4. Double-Blind Implementation Issues

In my role as program chair for several conferences, I implemented a double-blind reviewing process. Double-blind reviewing requires more work on the part of the program committee chair and program committee.

Software support and conflicts. The submission software should automate tracking and enforcing conflicts. Authors, committee members, and the program chair must enter conflicts. The program chair should review conflicts for consistency, and resolve any inconsistencies. Software automation can help ease this process. The submission software must ensure conflicted committee members never see reviews, rankings, or the reviewers of conflicting submissions (e.g., from their institution, collaborators, and current and former PhD students). The program chair must also ensure committee members leave the room during discussions of their conflict submissions. Reviewers and authors should specify individuals and institutions with whom they have a conflict. The authors should simply select from a list of potential reviewers and enter in a separate list their institutional and personal conflicts. The committee members should also provide such a list. Committee members could select conflicts from a list of submitting authors for the past several years. The reviewing software should persist and track conflicts over time (see Section 6).

At the beginning of the reviewing process, the program chair should remind the reviewers that knowing the authors' names and institutions before reading a submission can introduce positive and negative bias. The reviewers' current opinions and experiences (or lack thereof) with work from any individual should not influence the evaluation of the current submission. Reviewers should not endeavor to discover the authors, but should review any related work needed to determine the novelty of the submission.

Double-blind until accept. I suggest *double-blind until accept,* such that papers that are not accepted (the majority of papers at many conferences), will remain double-blinded when submitted elsewhere, where they may encounter some of the same reviewers or committee members. The argument for "light double-blind [9]" is that only your initial review need be blinded because a more pressing concern is that reviewers need to evaluate the submissions' novelty over the authors' own prior work. Unfortunately, there is no evidence in the scientific literature about when is best to unblind. This concern is however addressed (1) if authors cite their own prior work properly, as prescribed above; (2) if reviewers may appeal to the program chair if they believe double-blind is an obstacle to evaluating the submission (e.g., to determine access to special equipment or methods); and (3) when conflicts of interest are correct (see below). Authors who do not cite their own work properly should be penalized regardless.

Unblinding after the first review is submitted is insufficient because the reviewing process is far from over at this point. In the next phases, the other reviewers opinions are revealed. Are you willing to stand by your contrary judgement if the submission is from MIT or Nowhere University? Reviewer biases for and against authors may taint online discussions, the author response process, and the program committee meeting, all critical steps of the decisions making process.

Other mechanisms should protect against errors, such as conflict of interest mistakes. For example, the corresponding author may forget to enter conflicts for a co-author. The PC members must also list conflicts and the software should store and keep this information for each author historically. The duplication and

historical tracking of conflicts will help insure against conflict mistakes. If committee members ask colleagues for additional reviews, they must correspond with the program chair to avoid conflicts. No person with a conflict with the submission should see the reviews, reviewer names, ranking, or stay in the room during the discussion of the paper. This element is key to single-blind as well as double-blind reviewing and ensures the privacy of reviewers.

With these processes, author identifiers, and adding automated tool support for detecting conflicts, for example, mining the ACM Digital Library, I believe that errors in conflicts can be made very low.

5. External Review Committee

In addition to program committee reviews, many SIGPLAN conferences in the past obtained additional ad hoc *outside* (of the program committee) reviews. The goal of an outside review is to ensure a thoroughly expert review. With single-blind reviewing, the process of selecting ad hoc reviewers can be distributed among committee members. With double-blind reviewing, the same process is very error prone. For example, the program chairs of PLDI 2007, PLDI 2008, and ASPLOS 2006 took on this task, which consumed an enormous amount of time and email bandwidth.

The community seems to be now converging on using a version of a formal *external review committee* and/or tiered reviewing to solve problems of increased numbers of submission, generating expert reviews and handling double-blind conflicts. This approach was pioneered in SIGPLAN at ISMM 2008 by Steve Blackburn and is now common in many conferences, and especially large ones. The program chair selects the external review committee to complement and extend the expertise of the program committee.

The review committee has the same accountability as the program committee and follows all the same processes for conflicts, bidding on papers, reviewing standards, but they generally review fewer papers and do not attend the program committee meeting. A critical advantage of the external committee is handling the increasing number of submissions. The review committee should be sized to manage reviewer load for the expected number of submissions and reviews. They should add wide expertise to increase review quality and should only review submissions where they add an expert review. Depending on the number of submissions and reviewing rounds, external reviews could help with the first round of reviewing to generate 3 reviews per paper from the both committees, make a first cut, and then additional reviews from program committee members could be added only for these remaining papers. Program committee members would thus review higher quality submissions on average and ones more likely to be discussed at the committee meeting.

The program chair should apply the same submission assignment process, including conflict of interest procedures, to the external review and program committees. Compared to obtaining ad hoc external reviews for each submission, the review committee reduces the chance for errors and eases the burden on the program chair. Hopefully, it also improves reviewer quality because: (1) The review committee is transparent. (2) They are systematically selected to improve breadth and depth of expertise. (3) Since they review more than one paper, they can make relative judgments. (4) They can be made just as accountable by including them in the author response cycle.

Since external committee members do not have to travel, are selected, and acknowledged together with the program committee in the proceedings and on the web page, they have been willing to serve. This practice is now common in SIGPLAN and other conferences and is working well.

6. Conference Software and Resubmission

The conference software should (i) automate conflicts to reduce errors to a negligible number. With historical co-authorship information in a database, updated year, conflicts errors will be rare. (ii) The software should use information retrieval techniques to predict reviewer expertise to assist reviewers with bidding and program chairs with reviewer assignments. The software should mine prior reviewer

publications from the ACM DL and other sources for their expertise and score reviewers for expertise on each submission based on their publication history.

The software should retain prior submissions, their reviews, author responses, and a summary of differences with the new submission. I believe that this material should be made available to the current reviewers after their initial review. The VLDB community is experimenting with some policies in this regard. Because reviewing is hard work and consumes valuable researcher time, better mechanisms are needed encourage authors to incorporate prior comments, improve the work, and to utilize the prior reviews. An ACM and SIGPLAN wide policy is necessary to consistently give access to prior submissions and their reviews.

7. Author Response and Final Versions

In *author response*, also called *rebuttal*, reviewers enter their reviews, authors read the reviews, enter a response, and then reviewers make their final decisions. The purpose is to provide authors a forum to correct and directly address issues raised in the reviews. Unlike many journals, authors at this point have no opportunity to revise their submissions before reviewers make their decisions. However, reviewers may require revisions and authors are encouraged and always free to revise the final version of an accepted conference paper. All the above processes seek to improve reviewing quality, minimize some objections to double-blind reviewing, reduce errors, and ultimately improve the outcomes of the process. Even with these improvements however, scientific reviewing remains an imperfect human process that needs to continue to evolve.

I also recommend an author response phase to the reviewing process. I believe author response has the following benefits: (1) Because authors have the chance to correct reviews, reviewers are more accountable. (2) Reviewers get their reviews finished well in advance of the program committee meeting, which precludes reviewing in a rush on the airplane traveling to the meeting. (3) The program chair has time to obtain additional reviews if no reviewer is an expert or would like a specific type of review. The program chair should require reviewers to submit about a week or so before the program committee meeting. The authors should have two or three days in which to compose a response to the reviews, which includes answering reviewer questions, addressing concerns and issues raised in the reviews, and correcting any errors. The reviewers should read the responses and adjust their reviews accordingly before the meeting. At the meeting, the reviewer who leads the discussion should summarize the contents of the response. This practice is now common for many conferences.

Page limits on final versions seem counterproductive to encouraging research quality. Recent new policies in SIGARCH and SIGPLAN encourage inclusive references. Whereas in the past *submission page limits*, which help manage reviewer workload, applied to the entire submission, now several conferences now only apply page limits to the text and the bibliography has no page limit. This policy encourages appropriate references to closely related work. In the digital age, there seems to be no compelling reason to limit page counts for the archival article. Currently, authors can add archive additional material as an appendix in The ACM Digital Library, which requires no changes to conference publication policies.

7. Program Committee Submissions and Nepotism

Submitters have a conflict with other submitters. If this conflict were the only consideration, the best policy would forbid submission from reviewers and PLDI used to have this policy. However, Assistent Professors, in particular, were refusing to serve on the committee in order to submit. They rightly determined that publishing their research was paramount because their jobs depend on timely publication. However, they also benefit from serving on committees where they interact with senior researchers on how research is evaluated. Forbidding submissions from committee members also discouraged prolific senior researchers with graduate students from serving, or precluded submission of ready research, which disadvantaged their graduate students. This combination, the desire to encourage research service at all levels, and the desire to balance publishing needs and the reviewing needs of the community has led SIGPLAN (and other SIGS) to include committee submissions. The following best practices limit, but do not eliminate potential abuses.

To combat explicit bias against other submissions, the program chair should monitor the recommendations of all reviewers, and identify reviewers that are substantially more negative than other reviewers in general or an outlier on many submissions. After reading the reviews, since reviewers will vary in their judgements, the chair could encourage the reviewer to reconsider their judgements, discount the scores when choosing papers for discussion, and encourage online discussion of impacted papers.

To combat nepotism for program committee member submissions due to bias for students and colleagues other committee members, as found in related settings [8], the program chair should handle these submissions in a process before or after the program committee meeting. This eliminates conflicted authors leaving the room while their paper is discussed, revealing author and/or institutional identity. External review committees of PLDI, OOPSLA, and other conferences are now handling program committee submissions in a separate process, sometimes only with the external committee members to reduce the potential for nepotism for conferences with program committee submission.

9. In Person Program Committee Meetings

As far as I am aware, no research evaluates the impact of in-person meetings on publication quality compared to online discussion, an editor-only decision from reviews, or other decision processes. I believe that conference program committees play a critical role in training junior researchers, establishing shared research values, and creating inclusive and collegial communities that value variety in problem selection, approach, and methodologies.

Consider the Journal processes, in which the Editor-in-Chief and associate editors serve for multiple years, e.g., 3 to 6 years, and make all final decisions. Although many other researchers contribute reviews, reviewers may not even read the other reviews. This process concentrates publication power in relatively few senior researchers and does not promote open discussion of research evaluation. Although many journals publish standards, there is no transparency in how they are applied. Furthermore, many manuscripts are desk-rejected and never sent out for review. In contrast in the conference system, every submission is reviewed. Furthermore, program chairs and program committees turn over every year. This process means that a larger and more diverse set of researchers make publication decisions. At an in-person PC meeting, reviewer discussions are public to the rest of the committee and thus more transparent. I believe this more open process encourages reviewer accountability and likely leads to more consistent application of research criteria (e.g., novelty and methods) and explicit discussions of these standards at the meeting. However, applying consistent standards and evolving them to meet community needs from year to year have similar challenges as with journal reviewing. Ensuring that there is some overlap of the committee, but not too much, e.g., 10% to 15%, can provide year-to-year consistency.

Two additional important features of conference reviewing and in person meetings are community building and the training of recent PhDs. A non-negligible fraction of most SIGPLAN committees include recent PhDs, e.g., 10% to 30%. These meetings are often the first time recent PhDs meet and have substantial research interactions with a variety of senior researchers. Committee meetings are shared research experiences, where researchers learn about new problems, solutions, and methods, on a range of topics, not only for the submissions they review. While decision making is sometimes heated and factions may still form, factions are exposed and their power to dominate a discipline is reduced. The community benefits from this wider exposure to new ideas and research values. These benefits are especially critical to recent PhDs, who are in the process of defining an independent research program, usually for the first time.

10. Objections to Double-Blind

Although the majority of the SIGPLAN community is strongly in favor of double-blind reviewing, a vocal minority is not. Good processes and policies mitigate many of their objections. Snodgrass presents a list of objections and frequently asked questions about double-blind reviewing, which I recommend reading [5]. In my experience, the most objections come from *prolific researchers* (more senior researchers with many influential publications), who believe that it is ineffective to double-blind submissions and/or that it works against prolific authors. I discuss three common objections.

Some reviewers complain that it eliminates part of the benefit of program committee membership. For instance, SIGPLAN program committee work usually requires reading 15 to 30 papers. A benefit of this service work is gaining some global knowledge about the field. By removing authors information, the reviewer no longer learns who is doing which research, although the accepted work is revealed in the conference proceedings.

Some reviewers object to double-blind reviewing because they believe that they can identify authors based on the submission, even if authors follow the above guidelines. Research bears out that authors cite themselves more than other authors, and thus established, prolific researchers can often be identified through their citation list [1]. However, the very act of omitting author details on the paper has two distinct benefits. First, it reminds authors that they should endeavor not to reveal themselves through their citations or otherwise. Second, it reminds reviewers that they should judge the paper on its merits rather than based on whomever they guess the authors might be.

Some researchers believe that double-blind is mainly a tool to reduce positive bias *for* the submissions of prolific authors. For example, Schulzrinee writes "Double-blind reviewing primarily affects the perceived positive bias towards prolific, well-known authors" [13]. After SIGMOD adopted double-blind, Madden and DeWitt studied the impact of double-blind on prolific authors [11], showing that prolific authors acceptance rates were not influenced by double-blind [11]. A second study using the median instead of the mean on the same data reached the opposite conclusion [12]. The literature analysis by Snodgrass also reveals a conflicted reality [3]. Apparently, some reviewers actually hold prolific authors to higher standards or tire of their work, penalizing them, thus double-blind reviewing may help prolific authors. On the other hand, some reviewers do favor prolific authors and this benefit diminishes with double-blind.

The vast majority of the double-blind literature, however, is concerned with how double-blind affects racial and gender bias against historically discriminated groups. Unfortunately gender and racial biases are still prevalent, but the good news is that double-blind reduces bias and improves publication quality.

A more relevant study for computing and the SIGMOD data set would be to examine the gender and race of the authors of submissions and determine if double-blind changed the ratio of initial submissions (to evaluate perceived fairness) and their successes rates (to evaluate bias) from historically under-represented groups in computing. The under-representation of women, blacks, and hispanics in computer science compared to society as a whole, and their higher rate of attrition even for those who start out in computing, seems to have alluded the attention and priorities of those evaluating prolific authors in computing.

Double-blind is intended to help level the playing field for authors that *may or may not be prolific*, regardless of their race or gender, who have historically have been subject to explicit and implicit bias.

7. Conclusion

The goal of peer reviewing is to select publications that clearly present original ideas that move science forward in promising directions, use suitable evaluation methodologies, and make appropriate conclusions. Success at this task benefits researchers (prolific or otherwise), science, and the world. Double-blind reviewing improves the quality of decision making and the perception of fairness by increasing the focus of the evaluation process on the actual submission, rather than the authors. Can bad actors subvert double-blind? Unfortunately, yes. Is double-blind reviewing perfect? No, but double-blind reviewing improves fairness and quality, and all ACM and SIGPLAN conferences and journals should use it.

Notes

Although the majority of this material is original, some of it appeared in my 2008 SIGNOTICES article [14] and some was inspired by Snodgrass [3,4]. Of course, my personal experiences as a program committee member reviewer and program chair of ASPLOS, PACT, ISMM, CGO, and PLDI, informed this document. Discussions with colleagues, in particular with Sarita Adve, Steve Blackburn, Emery Berger, Mark Hill, and Margaret Martonosi, influenced my recommendations.

References

[1] S. Hill and F. Provost. The myth of the double-blind review?: Author identification using only citations, *ACM SIGKDD Explorations Newsletter*, 5(2):179--184, 2003.

[2] D. N. Laband and M. J. Piette. Citation analysis of blinded peer review. *The Journal of the American Medical Association (JAMA): The Second International Congress on Peer Review in Biomedical Publication*, 272(2):147--149, July 1994.

[3] R. T. Snodgrass. Single-versus double-blind reviewing: An analysis of the literature, *ACM SIGMOD Record*, 35(3):8--21, 2006.

[4] R. T. Snodgrass. Editorial: Single-versus double-blind reviewing reviewing, *ACM Transactions on Database Systems (TODS)*, 32(1):1--31, 2007.

[5] R. T. Snodgrass. Frequently asked questions about double-blind reviewing, *ACM SIGMOD Record*, 36(1):60--62, 2007.

[6] A. K. H. Tung. Impact of double-blind reviewing on SIGMOD publication: A more detailed analysis, *ACM SIGMOD Record*, 35(3):6--7, 2006.

[7] D. Watson, A. C. Andersen, and J. Hjorth. Mysterious disappearance of female investigators, *Nature*, 436(7048):174, July 2005.

[8] C. Wenneras and A. Wold, Nepotism and sexism in peer-review, *Nature*, 387(6 631):341--343, May 1997.

[9] M. Hicks, POPL'12 Program Chair's Report (or, how to run a medium-sized conference), January 2012.

[10] S. M. Blackburn, Survey results from authors of 2015 PLDI submissions, May 2015.

[11] S. Madden and DeWitt, "Impact of Double-Blind Reviewing on SIGMOD Publication Rates," ACM SIGMOD Record 35(2):29-32, June 2006.

[12] A. K. H. Tung, "Impact of Double Blind Reviewing on SIGMOD Publication: A More Detail Analysis," 2 pages, July 2006.

[13] H. Schulzrinne, "Double-Blind Reviewing --- More Placebo Than Miracle Cure?" ACM SIGCOMM Computer Communication Review archive, Volume 39 Issue 2, Pages 56-59, April 2009.

[14] K. S. McKinley, Editorial: Improving Publication Quality by Reducing Bias with Double-Blind Reviewing, External Review Committees, and Author Response, ACM SIGPLAN Notices, 43(8):5--9, August 2008.