

Doubly blind reviewing in IACR conferences, revisited

Ran Canetti
Boston University and Tel Aviv University

August 16, 2012

Abstract

We provide some background on the practice of doubly-blind reviewing (aka “anonymous submissions”) and discuss its advantages, disadvantages, and side effects. While much of the discussion is general, we concentrate on the specific case of IACR conferences. We conclude with an opinion and a suggested course of action. We hope that this white-paper will prompt open discussion within the IACR community on double blind reviewing in IACR conferences.

1 Background

Evaluating the merit of scientific work is an inherently subjective process. While evaluating basic aspects such as correctness and appropriate representation of the state of the art appear to be straightforward and objective, other aspects such as importance, scientific interest, level of novelty, clarity of exposition, or projected impact on industry and society, are highly elusive and lack objective criteria. In fact, even evaluating correctness and the representation of other work becomes subjective at times.

While it is often argued that subjectivity is an essential ingredient in the academic review process and key to scientific progress, it is also recognized that subjectivity carries within it the peril of having irrelevant data, biases and emotions interfere with the evaluation process. One potential cause for such extraneous considerations is information on authors, such as their identity, academic stature, seniority, institution, nationality, or gender. Indeed, while this information is largely perceived as extraneous to the academic merit of the work, it is believed to inadvertently affect reviews even when the reviewer intends to be unbiased and professional.

In order to mitigate this perceived deficiency of subjective academic review of scientific work, a number of academic venues have adopted a “doubly blind” review process, where the identities and affiliations of the authors are hidden from the reviewers (but not from the editors in journals or program chairs in conferences). While in some scientific disciplines double blind reviewing (DBR, also often called “anonymous submissions”) has been largely adopted, in others it has been contentious and in yet others it is almost non-existent. Specifically, in biology, chemistry and physics most prominent venues do not practice DBR (see e.g. the decision of Nature [9]). In Mathematics DBR is seldom practiced. In social sciences and economics it is more common, though still controversial. See e.g. [2] for a study of the effect of doubly blind reviewing in a prominent Economics journal.

Within Computer Science there is variance. While most theory oriented journals and conferences do not practice DBR, in other areas DBR is commonplace. For instance, most leading conferences in computer

vision practice DBR. In Networking, SIGCOMM practices DBR but several other leading venues do not. Most prominent venues in Databases do not practice DBR (with the exception of SIGMOD and TODS). Information Theory venues generally do not use DBR. In security-related venues not affiliated with IACR the situation is again a mix. For instance, ACM CCS and IEEE S&P practice DBR, while CSF does not.

It should be noted that a number of variants of DBR have been proposed and used, resulting in different properties and side effects. One parameter here is to what extent authors are expected to alter their behavior in order to preserve their anonymity. (E.g., is early online posting acceptable? Is referencing supporting material that violates anonymity acceptable?) Another parameter is whether author identities are to be disclosed to reviewers after the review process is done.

1.1 The case of IACR conferences

In the first Crypto and Eurocrypt conferences (1981 to 1988) author names and affiliations were included in the information accessible to the referees. In Crypto 1989, the PC chair Gilles Brassard instituted DBR. Since then, most IACR conferences have practiced DBR, with the support of the IACR Board. In the minutes of IACR board meetings, DBR is mentioned three times. The IACR Board meeting minutes at Crypto 1997 says:

Anonymous refereeing was introduced by Gilles Brassard when he was Program Chair of CRYPTO '89. After some discussion there was a sense that the Board was in favour of retaining anonymous refereeing.

The board meeting that approved the establishment of the eprint server (May 1999) notes:

It was ensured that authors are allowed to announce their results in public when they are in an anonymous refereeing process, that they can tell (and give away papers to) colleagues who work on similar matters and should know about an author's results. If an author announces a result widely, and committee members are on the distribution list, they should not be removed just because the paper is in submission. Authors are allowed to give talks on their papers and submit them to existing preprint servers, which will usually be announced widely. On the other hand, it is not intended that a submitter send letters to all the committee members saying who wrote which paper. Anonymous submission just means that papers are submitted without author's names and too obvious references.

The minutes of the IACR Board meeting at crypto 2006 (which was chaired by Cynthia Dwork and did not practice DBR) read:

Anonymous submissions were discussed next. Clark (via telephone) led the discussion, describing transactions related to the submission policy for EUROCRYPT 2007. A lengthy discussion ensued. A vote was called for mandatory anonymity for submissions (throughout the full evaluation process) to IACR conferences (CRYPTO, EUROCRYPT, and ASIACRYPT). 16 for, 0 against, 1 abstention.

That year, there was a membership ballot regarding whether to mandate anonymous submissions in IACR conferences or to leave the decision to the PC. The vote was 216 in favor of mandating anonymous submissions against 83 against, with 25 abstentions. It is noted that the decision applies only to IACR flagship conferences (Crypto, Eurocrypt and Asiacrypt). Still, all IACR-affiliated conferences practice DBR, except for TCC which does not.

With the exception of Section 8 in [1], we have not found additional records of public discussion for the rationale for the DBR practice in IACR conferences, either before or after the 2006 vote. Indeed, the lack of such documented debate is one of the main motivations for this white-paper.

2 Pros and Cons of double blind review

We list some of the common arguments for and against DBR. While most of the discussion is general, some points are specific to IACR conferences.

Pros. The main argument for DBR is that the best way to prevent extraneous information from interfering with the review process is to make sure the information is not available to the reviewer in the first place. This argument addresses two quite different cases: One in which the reviewer has all the best intentions to generate a professional and unbiased review, and one in which the reviewer knowingly intends to bias the review based on the extraneous information. Furthermore, the argument is that removing said information relieves reviewers from having to deal with the temptation to bias the review, and this helps reviewers (which are after all human) stay professional. A related argument is that even if the reviewer knows the identity of the authors with certainty or near certainty, the fact that the submission is anonymous makes it easier for the reviewer who wants to stay professional to put aside the temptation to bias the review based on the author. This is especially the case when the authors are close to the reviewer.

Indeed, studies show that having a journal move to DBR results in an increase in the number of citations to papers published in the journal [8]. This is interpreted as supporting evidence for an increase in quality (although number of citations is a rather superficial and misleading measure and is often not correlated with real scientific value).

Cons. Many quite different arguments against DBR have been made. We try to organize them into three levels. Of course, the partitioning is rough and some arguments can be viewed as spanning more than one level. Also, the arguments are certainly not listed in order of importance. Indeed, we find the arguments towards the end (the “side effects”) to be the strongest.

Effectiveness. One set of arguments challenges the effectiveness of DBR. Indeed, in the study of [2], about half of the time the reviewer could correctly guess the identities of the authors. [5] shows that when deploying automated tools the effectiveness of DBR is even lower. Certainly, using the “soft” form of DBR where early online posting of results is acceptable and commonly practiced, DBR is quite ineffective. However, given modern search tools preserving anonymity is hard even if early posting is disallowed, unless authors are extremely careful or virtually unknown.

Direct influence on review quality. A second class of arguments point to direct negative influence of DBR on the review process, regardless of its effectiveness. One argument is that DBR does not allow reviewers to link the submission with other papers (or submissions) by the same author or research group. In contrast, such linkage is often an important ingredient in the evaluation of the novelty and importance of scientific work. Indeed, as eloquently argued in [6, 7], scientific work is often best evaluated in the larger context of the research done by an individual or research group.

A related argument is that author names and reputations often play a useful role in the reviewers decisions regarding which aspects of a paper to study more closely (e.g., which proofs to double check) — and legitimately so. This is especially relevant in conference reviews, where program committee members need to make the overall best use of their very limited time. (An argument against this criticism is that determining correctness or plausibility based on author reputation is dangerous and may bias the reviews against unknown authors. Counter arguments would say that trustworthy reviewers would only use such information to manage their time and not let it affect their overall opinion, and that using such a reputation-based system does penalize carelessness and motivates careful verification of claims by authors.)

More generally, it has been argued that reviewers should be given full information regarding the evaluated work *as a principle*. Indeed, only trustworthy individuals should be given the task (or privilege) to review works of other in the first place. If an individual is not trustworthy, then DBR would certainly not help and the individual should not be reviewing papers at all. Furthermore, it is argued, trusting a reviewer with such sensitive information as author names may in fact help her put on the “reviewer robe” and transcend from egotistical considerations (see e.g. [4]).

Side effects. Next come some side effects (or, indirect consequences) in the DBR process:

Underhanded bias. One rather disturbing side effect is that of enabling “underhanded bias”: Reviewers who learn (or, guess) the identity of the authors of a paper in spite of the anonymity requirement might use the veil of anonymity to ‘kill’ or ‘push for’ the paper in a more ruthless and effective way than would be possible if names of authors were known (and thus allegiances would be obvious). The study in [2] shows that this is a real concern. In fact, the one most numerically significant result of that study is in the quantitative measure of this sort of bias (see Table 9 there).

To mitigate this side effect, some venues instantiate severe conflict-of-interest rules that prohibit researchers that have had any joint research with the authors in the past several years from serving as a reviewer. However, the effectiveness of such policy in mitigating underhanded bias is doubtful; furthermore, a significant negative side effect of this policy is that it rules out many of the more knowledgeable potential reviewers, thus reducing the overall professional level of the review process. this side effect is especially significant in highly collaborative and relatively small communities.

Post-review activity. It has been argued that the inability to link between authors and submissions often has negative effects even after the review process is over. Indeed, reviewers are expected to not use the ideas and results reported in rejected submissions in their own research. However, sometimes a reviewer might need to be able to link a rejected submission with its author — often for protecting the credit of the rejected work. For instance, a reviewer may be told similar or related results by a colleague, and might think that connecting the two sets of authors might preserve the credit due to the rejected work, or might advance joint research. Inhibiting such potential interactions may interfere with scientific progress. (It should be noted that this side effect vanishes if author identities are revealed to the reviewers as soon as the review process is over.)

Half-baked submissions. Yet another side effect, reported by many, is that DBR increases the number of half-baked and incomplete submissions. Indeed, it can be argued that the veil of anonymity and the lack of

“reputation system” increases authors’ incentive to submit works before they are complete. (This concern too can be mitigated by revealing the author identities to the reviewers once the review process is complete.)

Considering related public information. Another side effect is that, in order to comply with DBR requirements, authors are unable to include supporting information that is often very relevant to the merit of the submitted work. This includes data such as the results of public experiments, websites, code, use-cases, or public reviews that inevitably reveals the identity of the authors. This is a very real issue, especially in more practice-oriented submissions. Some venues (such as CVPR [3]) allow authors to upload anonymized supporting information. However, this is arguably only a partial solution since some information is inherently “live” and public.

Early publication. DBR that’s interpreted in a strict sense that does allow authors to post their results in the public domain during (and even before) their work is being reviewed can cause a delay of months and even years in the publication of results. In today’s environment where much of scientific work is published immediately in online unrefereed repositories, this delay has a significant negative effect on the impact of the delayed works and on scientific progress in general. Furthermore, questions of precedence and credit can (and often do!) result directly from such delays.

Constraining collaborative reviews. In many conferences (IACR conferences included) the paper review and program selection process is a highly collaborative and open one. Following an individual review phase, all PC members (except those with declared conflict of interest) contribute to the review and discussion on each submission. Often, papers are also discussed in groups and compared to each other. This open discussion is a crucial part of the review and often allows obtaining a much better understanding of the submissions and their contribution than simple concatenation of the original individual reviews.

The anonymity requirement is an impediment to the openness of this collaborative process. PC members cannot know which other PC members have declared conflict of interest (COI) with which papers — indeed, such information might well compromise the anonymity. Consequently, the discussion turns out to be much less collaborative, resulting in a worse quality of the overall review.¹

3 An opinion

At first glance (or gut feeling), DBR seems like a clean and desirable practice. It helps keeping the reviewer focused on the technical merit of the paper and does not inundate her with information that only generates “emotional noise”. Taking a more global view, it seems plausible that DBR would promote an open research community and suppress sectarianism (or “cliques” in research). Even when not completely effective, DBR highlights the expectation to keep author identity outside the equation.

¹A good support tool can somewhat help here. Indeed, Shai Halevi’s tool, which is currently used in IACR conferences, helps in creating individual review pages which only non-COId PC members can access. However, this is still a very partial solution. For one, the current software does not address discussing papers in groups or serendipitous consultations. More importantly, even a perfect tool cannot resolve the inherent contradiction between the need to preserve author anonymity and the need for open collaborative review that respects COI limitations.

A second look brings up a number of serious deficiencies. A first such deficiency is the inherent contradiction between effective DBR and the need to allow quick unrefereed posting of research results: The two cannot go hand in hand, neither in principle nor in practice. Here IACR took the path of allowing quick unrefereed posting. While this path is clearly the right one for supporting scientific progress, it renders DBR a rather lame practice with very selective effectiveness.

Another deficiency that's very disturbing is the introduction of underhanded bias, as discussed in the previous section. The fact that DBR is only selectively effective due to the eprint archive and other quick publication venues makes this side-effect much worse.

To this, one should add the formal inability to include public supporting materials in submissions, the inhibitive effect on the openness of the collaborative review process, and the administrative overhead.²

We conclude that, in spite of its theoretical appeal, DBR (at least in its current form in IACR conferences) is a bad compromise that on the one hand is not really effective, and on the other hand still causes a number of serious negative side effects. More generally, there seems to be an inherent contradiction between DBR and scientific progress.

Furthermore, the situation today is different from that in 2006, and certainly 1989. The pace of new research results in crypto (and likewise the use of eprint and other immediate publication venues) has increased. The effectiveness of DBR in today's searchable environment has diminished. On the other hand, the overhead of DBR has increased: PCs of IACR conferences have considerably grown in size, and that a considerable number of submissions are authored by PC members. PC members and especially PC chairs have to constantly worry about finding creative compromises between the requirement to preserve anonymity of submissions and the desire to enable effective review.

So it is time to revisit the 2006 vote. But what method is preferred? Several alternatives have been suggested, including voluntary anonymity, disclosing author names after the individual review stage, disclosing author names after tentative decisions have been made. However, while these alternatives might make for interesting experiments, neither of them appears to mitigate all the drawbacks mentioned above.

At the same time we believe that the IACR community has matured and can better handle an open review process that makes author information available to all PC members, along with strict expectation to keep the evaluation of the merit of submissions unbiased, unsectarian, and professional. So my vote would be to return to full disclosure of names of authors of submissions.

References

- [1] E. Biham. Updating the IACR Publication Infrastructure by a Preprint Server. <http://eprint.iacr.org/misc/pubrep.pdf>
- [2] Rebecca M. Blank. The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from The American Economic Review. The American Economic Review Vol. 81, No. 5

²As an personal point of reference, I have chaired TCC'08 and co-chaired Crypto'12. The first did not use DBR whereas the second did. I did not notice any more incidents of (or attempts to) bias or sectarianism in TCC'08 than in Crypto'12. On the other hand, the review process was considerably more open and collaborative in TCC'08 - largely, in my eyes, due to the lack of DBR constraints. (I'd like to stress that I do not in any way think that the Crypto'12 PC was any less professional or dedicated than the TCC'08 PC. I attribute the difference exclusively to the DBR constraints.)

(Dec., 1991), pp. 1041-1067 Published by: American Economic Association Article Stable URL:
<http://www.jstor.org/stable/2006906>

- [3] Computer Vision and Pattern Recognition (CVPR) conference. <http://www.cvpr2012.org/program-details/papers/submission/author-guidelines>
- [4] Oded Goldreich. Critique of some trends in the TCS community. <http://www.wisdom.weizmann.ac.il/oded/toc-crit.html>
- [5] Shawndra Hill. The myth of the double-blind review: author identification using only citations. ACM SIGKDD Explorations Newsletter. Volume 5 Issue 2, December 2003 Pages 179-184.
- [6] Daniel Lemire. The case against double-blind peer review. <http://lemire.me/blog/archives/2011/04/28/the-case-against-double-blind-peer-review/>
- [7] Daniel Lemire. Is science more art or industry. <http://lemire.me/blog/archives/2011/04/29/is-science-more-art-or-industry/>
- [8] Kathryn S McKinley. Editorial: Improving Publication Quality by Reducing Bias with Double-Blind Reviewing and Author Response. ACM SIGPLAN Notices, 43(8):5–9, August 2008. <http://www.cs.utexas.edu/users/mckinley/notes/blind.html>
- [9] <http://www.nature.com/nature/journal/v451/n7179/full/451605b.html>