Editorial: Single- Versus Double-Blind Reviewing

Richard T. Snodgrass
University of Arizona

This editorial analyzes from a variety of perspectives the controversial issue of single-blind versus double-blind reviewing. In single-blind reviewing, the reviewer is unknown to the author, but the identity of the author is known to the reviewer. Double-blind reviewing is more symmetric: the identity of the author and the reviewer are not revealed to each other. We first examine the significant scholarly literature regarding blind reviewing. We then list six benefits claimed for double-blind reviewing and 21 possible costs. To compare these benefits and costs, we propose a double-blind policy for *TODS* that attempts to minimize the costs while retaining the core benefit of fairness that double-blind reviewing provides, and evaluate that policy against each of the listed benefits and costs. Following that is a general discussion considering several questions: what this has to do with *TODS*, does bias exist in computer science, and what is the appropriate decision procedure? We explore the "knobs" a policy design can manipulate to fine-tune a double-blind review policy. This editorial ends with a specific decision.

Categories and Subject Descriptors: H.2 [**Database Management**]: General

Additional Key Words and Phrases: anonymous citation, blinding efficacy, double-blind review, gender bias, single-blind review, status bias

1. INTRODUCTION

The peer review process is generally acknowledged as central to the advancement of scholarly knowledge. It is also vital to the advancement of individual careers.

With so much at stake, it is important to examine, and re-examine, issues pertaining to review quality on an on-going basis. Thus it is appropriate that controversy has arisen in our field pertaining to the practice of double-blind reviewing. "As scientists, we should rather welcome all occasions to reflect on the act of writing, evaluating, editing and publishing research findings. The issue of double-blind refereeing, which recurs periodically in scientific circles, provides us with such an opportunity" [Genest 1993, page 324].

Most database journals employ single-blind reviewing, in which the reviewer is unknown to the author, but the identity of the author is known to the reviewer. Others employ double-blind reviewing, in which the identity of the author and the reviewer are not revealed to each other. The arguments for double-blind reviewing

Author's address: Richard T. Snodgrass, Department of Computer Science, University of Arizona, Tucson, AZ 85721-0077, rts@cs.arizona.edu.
This is a preliminary release of an article accepted by ACM Transactions on Database Systems. The definitive version is currently in production at ACM and, when released, will supersede this version.

are that it is more fair and that it produces higher quality reviews. The arguments advanced against double-blind reviewing include that it has little effect, that it makes it more difficult for reviewers to comprehensively judge the paper, and that it is onerous to administrate [Ceci & Peters 1984].

The editorial first examines the now substantial scholarly literature regarding blind reviewing. This literature includes empirical studies from biomedicine, communication, computer science, economics, education, medicine, public health, physics, and psychology, retrospective analyses from computer science, ecology, economics, and medicine, and a quantitative meta-analysis from psychology. It is useful and instructive to learn what other disciplines, using diverse approaches, have discovered about blind reviewing.

Later sections provide a comparison of costs and benefits, a proposal for a double-blind reviewing procedure, and an general discussion of relevant questions. We end with a specific decision.

2. AN ANALYSIS OF THE LITERATURE

In the following[1], we first define the various terms used in the literature. We then examine in some depth the general issues of fairness. We end with a brief summary of this complex sociological question.

2.1 Terminology

ACM defines a *refereed journal* or *refereed conference* as one that "is subjected to a detailed peer review, following a defined, formal process according to a uniform set of criteria and standards." [2] Material appearing in such venues is distinguished from *formally reviewed material* ("subjected to a structured evaluation and critique procedure following a defined process uniformly applied as with refereeing, only without requiring that the tests of scholarly originality, novelty and importance be applied"), *reviewed* ("subjected to a more informal and not necessarily uniform process of volunteer review, with standards dependent upon the publication and the type of material"), *highly edited* ("professionally edited, usually by paid staff, with primary emphasis on exposition, graphic presentation, and editorial style rather than on content and substance"), and *unreviewed* ("published as submitted, with or without copyediting"). "Reviewing" in the present document refers to peer review for a refereed journal or conference.

Peer review is the use of predetermined reviewers, in the case of program committees, or ad hoc reviewers, in the case of reviewers for most journals, who individually read the submitted manuscript and prepare a written review. Sometimes, as in the case of some conference program committees, reviewers will subsequently either physically or electronically meet to discuss the papers to arrive at an editorial decision. For most journals, the Associate Editor handling the paper or the Editor-in-Chief will make the final editorial decision.

In the vast majority of refereed database conferences and journals, the identity of the reviewer(s) is not revealed to the author(s), ostensibly to ensure more objec-

[1]This section contains a subset of the literature analysis that appeared in a companion paper [Snodgrass 2006], with an added discussion of psychological studies of bias.
[2]ACM Policy on Pre-Publication Evaluation, at `http://www.acm.org/pubs/prepub_eval.html`

tive reviewing. This is termed *single-blind reviewing* or, less frequently, *one-eyed review* [Rosenblatt & Kirk 1980]. (The terms *reviewer* and *referee* are used interchangeably in the literature.)

In an effort to achieve more objective reviewing, a venue can also request that the identity of the author be removed from the submitted manuscript, a process termed *blinding the manuscript*. When the identity of the authors and their institutions is kept from the reviewers, this is termed double-blind reviewing. Note that the Editor-in-Chief and Program Chair, and generally the Associate Editor, are made aware of this information via a separate cover sheet not shared with the reviewers.

The psychological sciences utilize a different terminology that conveys a subtle philosophical shift. When the identities of the authors and reviewers are not revealed to each other, it is termed in these sciences a *masked* reviewing process. Note the symmetry of this terminology. The American Psychological Association *Guide to Preparing Manuscripts for Journal Publication* [Calfee & Valencia 2006] states, "Peer review is the backbone of the review process. Most APA journals, like the majority of other professional publications, practice anonymous, or masked, reviews. Authors and reviewers are unaware of each other's identities in most instances, an arrangement designed to make the process more impartial." The implication is that revealing either the reviewer's identity or the author's identity breaks the mask. Presumably single-blind reviewing would then be termed "non-masked," but the APA doesn't use the term. (The term *unmasking* denotes revealing the identity of a reviewer to a co-reviewer [van Rooyen 1999]; we don't consider that practice here.)

This editorial will use the terms single-blind and double-blind reviewing, as well as their respective three-letter acronyms, *SBR* and *DBR*.

Venues differ in who does the blinding/masking of a submission. We will use the term *author masking* when the author removes identification from the paper before submitting it and *editorial masking* when such identification (generally, author name and affiliation) is removed in the editorial process before sending the manuscript to the reviewers. Procedures differ in how aggressive is the required author masking and the actual editorial masking. Self-citations and other first-person references in the body of papers are generally retained in editorial masking. While author masking can be more thorough, because authors would know what kind of information is revealing, authors through various devious means can circumvent both kinds of masking.

Previous literature reviews (see the companion paper [Snodgrass 2006] for a survey) emphasized three primary aspects relevant to blind reviewing: fairness to authors (to unknown authors or to authors affiliated with unknown institutions, to less-published or to proficient authors, to both genders), review quality, and blinding efficacy. The following section addresses fairness; the companion paper contains a comprehensive survey of the other two aspects.

2.2 Fairness

The fundamental argument for double-blind reviewing is that it is more fair to authors (and thus, indirectly, to readers). The argument proceeds as follows: The judgment of whether a paper should be accepted for publication should be made on the basis of the paper alone: is what the submission states correct, insightful,

and an advancement of the state-of-the-art? The editorial judgment should not be made on extenuating circumstances such as who wrote the paper or the professional affiliations of the authors. By blinding the submission, the reviewers cannot take these peripheral aspects, which are not relevant, into account in their review.

Psychology researchers have studied bias in judgment. Expectations are one source of bias. Quality is judged relative to expectations, which is why a $6 cup of coffee tastes better than a $2 cup of coffee, holding everything else constant. This sort of bias is both conscious and unconscious. People know that they expect more expensive coffee to taste better, but they don't realize that the mere fact of having that expectation actually makes it taste better [Shiv et al. 2005]. Some reviewers have an expectation that "papers from ... are never any good."

There is now also a good amount of evidence from psychological studies that judgment often operates in an unconscious fashion [Greenwald & Banaji 1995]. People who don't *know* that they are biased tend to be those who are the most biased [Wilson & Brekke 1994]. And paradoxically, under some conditions, experience and motivation, both of which are in ample supply in reviewing, can accentuate some forms of judgment bias [Kardes et al. 2005]: "people believe that they learn a lot from experience even when experience is actually irrelevant" [ibid, pp. 149–150]. Such experience could involve the identity or affiliation of the author.

The relevant question then is whether these conscious and unconscious sources of bias actually come into play in reviewing of conference or journal articles. The analysis of fairness in the extant literature concerns (a) fairness to unknown authors or institutions, (b) fairness to prolific or to less-published authors, and (c) gender equity. There is also the related issue of the perception of fairness. The following sections will elaborate on each of these concerns.

2.2.1 *Fairness to Unknown Authors or Institutions.* Some evidence from retrospective and experimental studies suggest that when the authors' names and affiliations are known, reviewers may be biased against papers from unknown authors or institutions, termed *status bias* [Cox et al. 1993]. We now examine the studies that attempt to detect status bias, in chronological order.

A retrospective study of manuscripts that had been submitted to *The Physical Review* between 1948 and 1956 found that "some 91 per cent. of the papers by physicists in the foremost departments were accepted as against 72 per cent. from other universities" [Zuckerman & Merton 1971, page 85]. Two possible explanations were offered: status bias and "differences in the scientific quality of the manuscripts coming from different sources" [ibid]. Crane [1967] provided an alternative explanation, "that common viewpoints rather than personal ties could explain acceptance" [Dalton 1995, page 227], yet such common viewpoints would have been present in both SBR and DBR.

An early experiment found that "the effect of institutional prestige failed to attain significance in any one of the measures" [Mahoney et al. 1978, page 70]. "Experimental manuscripts were sent to 68 volunteer reviewers from two behavioristic journals. ... Institutional affiliation was also manipulated on the experimental manuscripts, with half allegedly emanating from a prestigious university or a relatively unknown college" [ibid].

Another retrospective study, this of the records of reviews of a society which

publishes research journals in two areas of the physical sciences, found large differences in how papers from minor and major universities are reviewed: "minor university authors are more frequently evaluated favourably (ie less critically) by minor university referees, while major university authors are more often evaluated favourably by major university referees than they are by those affiliated to minor universities. It would therefore appear that when referees and authors in these areas of the physical sciences share membership of national or institutional groups, the chances that the referees will be less critical are increased. ... Personal ties and extra-scientific preferences and prejudices might, of course, be playing a part as well. But it appears that, even in the absence of these personal factors, the scientific predispositions of referees still bias them towards less critical evaluation of colleagues who come from similar institutional or national groups, and so share to a greater extent sets of beliefs on what constitutes good research" [Gordon 1980, pp. 274–5].

A seminal experiment [Blank 1991] demonstrated status bias in reviewing more directly. In this experiment, every other paper that arrived at the *American Economic Review* was designated as double-blind. For these papers, an editorial assistant removed the name and affiliation of the author from the title page and typically scanned the first page for additional titles or notes that would identify the author (that is, those manuscripts were editorially masked). This experiment lasted for two years.

The relevant issue was "whether the ratio of acceptance rates between institutional ranks in the blind sample differs from the corresponding ratio in the nonblind sample." [Blank 1991, page 1053–1054]. It was found that this ratio did not differ for those at top-ranked departments and those at colleges and low-ranked universities. All other groups, in that important gray area where editorial judgment is most needed, had substantially lower acceptance rates in the blind sample than in the nonblind sample; in some cases, the acceptance rate dropped by more than 7 percentage points. She found similar differences with referee ratings between SBR and DBR.

A retrospective study of single-blind reviews for the *Journal of Pediatrics* and published in *JAMA* found only partial evidence for status bias, that "for the 147 brief reports, lower institutional rank was associated with lower rates of recommendation for acceptance by reviewers ($P < .001$). ... For the 258 major papers, however, there was no significant relationship between institutional rank and either the reviewer's recommendations (P=.409) or the acceptance rate (P=.508)" [Garfunkel et al. 1994, page 138].

Another retrospective analysis of single-blind reviews also published in *JAMA* found evidence of status bias at a coarse geographical level [Link 1998]. In this analysis of original research articles submitted to *Gastroenterology* during 1995 and 1996, it was found that "reviewers from the United States and outside the United States evaluate non-US papers similarly and evaluate papers submitted by US authors more favorably, with US reviewers having a significant preference for US papers" [Link 1998, page 246].

The experimental evidence is mixed concerning status bias present for top-ranked authors and institutions. The evidence is quite compelling that status bias is possi-

ble, perhaps prevalent, in SBR for most other authors and institutions, presumably for those papers most needing the critical evaluation of reviewers.

2.2.2 *Fairness to Prolific Authors.* There have been several studies that have looked at the impact of blinding on prolific authors; these studies were discussed in the companion paper [Snodgrass 2006]. Contradictory results from these studies render it impossible to say anything definitive about the impact of blinding on prolific authors. However, there does seem to be evidence of some kinds of bias with SBR.

2.2.3 *Gender Equity.* When reviewers know the identity of the author(s) of the submitted manuscript, gender bias is also a possibility. Several disciplines have launched in-depth studies based on concerns of gender equity.

Blank's experiment, described earlier, was in fact initiated due to concerns of gender bias. The *American Economic Review* journal had employed SBR for most of its recent history, except during a period of 1973–1979 when the then-current editor adopted DBR [Borts 1974]. In the mid-1980's, the American Economic Association's Committee on the Status of Women in the Economics Profession formally expressed its concern about "the potential negative effect on women's acceptance rates of a single-blind system" [Blank 1991, page 1045]. As a result, Blank was asked by the current editor of the *AER* and the Board of Editors to design and run a randomized experiment looking into this potential effect.

Due to the careful randomization design of this experiment, one can compare acceptance rates between the blind and nonblind samples, and indeed, there were striking differences. "For women, there is no significant difference in acceptance rates between the two samples. For men, acceptance rates are significantly higher in the nonblind sample." [ibid, page 1053]. When reviewers knew that that paper was authored by a male, they accepted a higher percentage (15%, versus 11%) than if the paper was blinded. "One can compare acceptance rates between the blind and nonblind samples without other control variables because the randomization process guarantees that papers by women (and men) in each sample have identical distributions of characteristics" [ibid].

Blank emphasized the core issue: "whether the *ratio* of male to female acceptance rates in the nonblind sample is different from that in the blind sample. In both samples, women's acceptance rates are lower than men's, but the differential in the blind sample is smaller. While women in the blind sample have an acceptance rate only 1 percentage point below that of men, their rate is 3.8 percentage points lower in the nonblind sample" [ibid]. Here the results were statistically insignificant, perhaps because there were too few observations of papers authored by women.

Would DBR result in a large increase in acceptances of papers by women? "While there is some indication in these data that women do slightly better under a double-blind system, both in terms of acceptance rates and referee ratings, these effects are relatively small and statistically insignificant. Thus, this paper provides little evidence that moving to a double-blind reviewing system will substantially increase the acceptance rate for papers by female economists" [ibid, page 1063]. Interestingly, the *American Economic Review* reverted to DBR after Blank' study was published.

The Modern Language Association's (MLA) experience was striking: going to DBR resulted in a large increase in acceptances by female authors. "Contributed papers at MLA meetings had first to survive a review stage before acceptance to be read. Prior to 1974, these papers were refereed with the author's name intact. In 1974, double-blind refereeing was tried with the effect that the number of women and of new investigators having papers accepted doubled from previous years. This number doubled again when repeated in 1975, until, by 1978, the proportion of acceptances among women and new researchers was comparable to that for men. The MLA Board subsequently decided in 1979 to use double-blind refereeing for all their publications" [Billard 1993, page 321]. The impetus for this change was the perception of gender bias. "A number of women complained to the Modern Language Association in the United States that there were surprisingly few articles by women in the association's journal, compared to what would be expected from the number of women members. It was suggested that the review processes were biased. The association vigorously denied this but under pressure instituted a blind reviewing procedure under which the names of the authors and their institutional affiliations were omitted from the material sent to the reviewer. The result was unequivocal: There was a dramatic rise in the acceptance of papers by female authors" [Horrobin 1982, page 217].

It is possible that the small observed effect in Blank's study (in contrast to the MLA experience) was due to the low number of submissions by women to *AER*. Certainly the computer science field is closer to economics than modern languages in its participation of women.

These studies show that revealing author identity, specifically the gender of the author, can sometimes have an effect on acceptance rates.

2.2.4 *The Perception of Fairness.* A *perception* of possible bias may be just as damaging as actual bias.

The Institute of Mathematical Statistics (IMS) New Researchers' Committee (NRC) report stated, "The NRC feels that the current system [SBR] has the potential for bias or perceived bias against NRs [new researchers], women and identifiable minorities, (a disproportionate number of the latter two categories are NRs)" [Altman et al. 1991, page 165]. In a response to discussants of that report, the NRC reasserted a year later, that "much of the value of double-blind refereeing lies in the community perception of fairness" [Altman et al. 1992, page 266].

The experience with this controversy at the IMS indicated a split between new researchers, which "strongly endorses double-blind refereeing. ... It seems likely that [this] represents the majority opinion among new researchers, although support for double-blind refereeing is not unanimous among new researchers" and senior members: " 'negative but sympathetic' ... seems to be a majority view among those senior enough to have been involved in the editing process" [Cox et al. 1993, page 311]. However, a survey to IMS members "indicates strong support for double-blind refereeing in the IMS journals" [ibid].

A responder to the IMS report [Cox et al. 1993] stated, "Refereeing is *perceived* by many writers as being subject to various kinds of biases: biases in favor of male or female, young or established, national or foreign researchers, working at small or large institutions, in well-developed or developing countries and so on.

Whether such biases are sufficiently strong and widespread to distort the whole review process is beyond the point. So long as the *potential* for abuse is there, we should guard against it, and double-blind refereeing is but one means of ensuring such protection" [Genest 1993, page 324] (emphasis in original).

2.3 Quality of Reviews

From the analysis in the companion paper [Snodgrass 2006], one must conclude that the jury is still out. It has not been shown convincingly that either SBR or DBR can, by revealing or by hiding the identity of the author and institution, increase the quality of the reviews of a submitted manuscript.

2.4 Efficacy of Blinding

The companion paper [Snodgrass 2006] examined a number of studies of blinding efficacy. Blank's conclusions apply to the many studies generally. "On the one hand, a substantial fraction—almost half—of the blind papers in this experiment could be identified by the referee. This indicates the extent to which no reviewing system can ever be fully anonymous. On the other hand, more than half of the papers in the blind sample *were* completely anonymous. A substantial fraction of submitted papers are not readily identified by reviewers in the field. ... Those blind papers that are correctly identified by the referees ... are skewed in favor of authors who are better known or who belong to networks that distribute their working papers more widely" [Blank 1991, pp. 1051–2]. Ceci and Peters conclude that "Although there are occasional lapses in the preparation of manuscripts by authors and failures to screen manuscripts by editorial staff, we are impressed by the overall efficiency of blind review" [Ceci & Peters 1984, page 1494].

2.5 Summary

"There is a long tradition attached to the peer review system. As *users* of science, we all depend on it: our professional realizations are based upon the work of others, and we count on journal (and book) editors to separate the wheat from the tares. Although there is no such thing as perfection, it would be a disservice to the profession if too many scientific writings addressed irrelevant issues or contained gross factual errors. As *producers* of science, it is also in our interest that the system be fair: favoritism, discrimination and condescension bring discredit on the entire operation and ultimately work against the discipline, even if individual benefits occasionally may accrue in the short term" [Genest 1993, page 324].

We have attempted here to summarize the many studies of the varied aspects of blind reviewing within a large number of disciplines.

Concerning the central issue of fairness, Blank's summary in 1991 of the literature still holds true fifteen years later. "In summary, the literature on single-blind versus double-blind reviewing spans a wide variety of disciplines and provides rather mixed results. Few of the empirical tabulations provide convincing evidence on the effects or non-effects of refereeing practices, largely because of their inability to control for other factors in the data. If not fully convincing, however, there is at least a disturbing amount of evidence in these studies that is consistent with the hypothesis of referee bias in single-blind reviewing" [page 1045]. Many studies provide evidence that DBR is more fair to authors from less-prestigious institutions

and to women authors. Such differences are likely to matter even more for highly-selective conferences and journals.

A companion paper [Snodgrass 2006] looked at related issues. Concerning quality of reviews, it is not known definitely whether either SBR or DBR results in a higher quality of reviews. Most of the studies discussed here and in the companion paper utilize editorial blinding, which has been shown to be successful about 60% of the time, across many disciplines. Removing text allusions and self-citations would increase success rates to perhaps 75%. The prevalence of DBR has increased dramatically over the last fifteen years, to the point where most scientific journals now employ double-blind reviewing.

3. AN ANALYSIS OF COSTS VERSUS BENEFITS

As the previous literature survey emphasized, there are demonstrable, albeit intangible *benefits* of a double-blind reviewing process. And it emphasized as well that there are very real and tangible *costs*, to the journal as well as to the author and perhaps to the reviewer, of adopting DBR. These costs depend in part on how the blinding is done, with the efficacy directly related to the effort expended. These costs also vary depending on the needs of the scholarly discipline covered by the journal in question and on the culture of the scholarly community served by that journal.

Journals strongly desire to fairly evaluate submitted manuscripts, while simultaneously keeping costs in control. The policy question before each journal and each scholarly publisher is thus the following. Is the documented benefit of equity worth the administrative cost? At what price fairness?

We now enumerate the benefits and costs of adopting DBR and proposes a specific procedure for *TODS* that attempts to minimize those costs while retaining the documented benefits.

3.1 Perceived Benefits of DBR

We first list benefits claimed for DBR.

DBR avoids a bias towards top-ranked authors and institutions. Perhaps SBR encourages an unconscious bias towards prominent authors or institutions.

DBR avoids a bias towards prolific authors. Prolific authors will be well-known, and may be treated more gently by reviewers.

DBR encourages higher-quality reviews. Perhaps by not knowing the identity of authors or their institutions, reviewers will write more authoritative reviews.

DBR is more fair to women authors. Studies have shown that revealing the gender of the author can have an effect on acceptance rates.

DBR is more fair to non-top-ranked authors and institutions. The evidence is quite compelling that status bias is possible, perhaps prevalent, in SBR for such authors.

There is a perception by some that DBR is more fair. Several task forces and committees have called for DBR.

3.2 Perceived Costs of DBR

There are very real costs to DBR that must be objectively considered when contemplating a change in a journal's editorial masking policy. Some of these costs apply

to most scholarly journals; other costs are specific to a computer science journal.

Quality of reviews is reduced. Perhaps DBR, by depriving the reviewer of the knowledge of the identity of the author(s), their institution, their prior work, or the biases identifiable through their prior work, reduces the quality of reviews.

Intentional bias is reduced. Perhaps DBR disallows intentional bias that attempts to counter unconscious bias by other reviewers.

It is difficult to mask submitted manuscripts. Editorial masking requires either blackening out identifying information from a paper (which is difficult if the paper is submitted electronically) or requires that the source be also submitted so that it can be edited. Author masking requires effort on the part of the author to remove identifying information.

The web can be searched by reviewers to find related papers (especially those that are similarly titled by the same author, such as technical reports and related papers), thus making it fairly easy to narrow down the set of suspected authors.

DBR increases the probability of non-novel papers getting accepted. Perhaps DBR will discourage reviewers from looking up related work, and so may artificially inflate the perceived contribution of the work.

It is difficult to fully mask systems papers. A mode of inquiry common in computer science is the creation of often extensive software artifacts, often over a period of several years. In this research methodology, a succession of papers is produced, each of which builds on the artifacts and prior papers. Thus it is difficult to impossible to fully mask manuscripts coming out of such projects, as the contributions are necessarily incremental.

DBR jeopardizes "power projects". There is a history of high-impact projects in computer science; examples include the Berkeley RAID project, the Cornell Program Synthesizer, Berkeley Ingres, IBM's System R, and UCSD Pascal. Such projects may not be compatible with DBR. Papers from these projects benefit from the "brand" that has been created by prior distribution of papers and software artifacts such as freely distributed source code; DBR does not allow this brand to be exploited or even acknowledged by the submitted manuscript.

Forward-looking systems work, in particular, suffers from anonymization. Such systems work can get obfuscated if the author tries to truly anonymize it, or it is really not anonymized.

Anonymization is a distraction to students learning to write clearly. Do we want to emphasize to students how to make something (in this case, authorship) unclear?

Some journal policies preclude DBR. As one example, *TODS* requires that submissions invited from conferences utilize some of the original reviewers from that conference's program committee in the review of the journal submission. DBR would preclude inviting papers in the future from conferences.

DBR precludes expanding a conference paper into a journal submission. A submission that is an expanded version of a conference paper, containing much of the same prose and often a similar if not identical title as the conference paper, cannot be easily masked.

DBR makes it difficult for conflicts of interest to be detected. How can a reviewer determine whether a conflict of interest exists with an author of a blinded manuscript?

DBR prevents the dissemination of research results before submission. Strictures on prior submission in DBR policies could prevent working papers and technical reports from being distributed for comment, because such dissemination could reduce the efficacy of author masking.

DBR gets papers rejected for failing to sufficiently cite the author's own work. Citations removed during the blinding of a submitted manuscript may cause a paper to be rejected because it does not cite relevant work.

DBR increases the possibility of plagiarism. With DBR, a reviewer might assume material is from one author and judge it legitimate, even if the material was actually plagiarized from another author.

DBR raises the possibility of self-plagiarism. With DBR, other work by the authors may be anonymously cited or not cited at all, making it more difficult for reviewers to catch self-plagiarism.

DBR places extra burdens on reviewers. DBR, through the anonymization process, may render the paper harder to understand or to review.

DBR decreases enthusiasm of reviewers. Perhaps DBR (and even the discussion about DBR) sends the message that the journal does not trust its reviewers.

DBR reduces timeliness. If the enthusiasm of reviewers is reduced, or if the process is cumbersome, timeliness may be adversely affected.

DBR places extra burdens on the Editor-in-Chief and Associate Editors. Depending on how the mechanism for ensuring DBR is designed, this mechanism may impose more work on the Editor-in-Chief (EiC) or on the Associate Editors (AEs).

DBR places extra burdens on the authors. Similarly, depending on how the mechanism for ensuring DBR is designed, this mechanism may impose more work on the authors.

SBR versus DBR should not be a local decision. There is the "cost" of incompatibility between the editorial policies of various journals serving a scholarly community. It is less confusing if these journals present a consistent set of policies. Otherwise a paper will have to be altered if it is rejected from one journal and subsequently submitted to another journal.

3.3 A Proposal for a Double-Blind Policy

To compare the costs and benefits of DBR, we propose a double-blind policy for *TODS* that attempts to minimize the costs while retaining the benefit of fairness.[3]

This proposal roughly follows the approach utilized by the ACM *SIGMOD* conference, which has been using DBR since 2001. That conference's double-blind policy specifics and procedures have been refined over the past six years, and have been applied to several thousand submissions; this policy is now well-known in the database community. An informal check of several other conferences and journals utilizing DBR indicated that the instructions from the *SIGMOD* conference are by far the most specific and helpful in their guidance to authors.

[3]This proposal was hammered out with the *TODS* Editorial Board over an intensive two-month discussion via email.

For each aspect (applicability, submission, reviewing, revision, acceptance, transition), we state the proposed policy, then review the rationale behind that proposal. The following section revisits the benefits and costs of adopting this policy.

As mentioned, two kinds of masking have been discussed in the prior literature: editorial blinding and author blinding. In general, there is a wide spectrum between the extremes of a manuscript that explicitly states its authorship (unmasked) and one for which there is no possibility of a reviewer discovering the identity of any of the authors (complete masking). Indeed, the latter situation is not feasible. There is always the potential of a reviewer discovering, through a web search, through conversation with colleagues, through knowledge of the literature, or through other means, the identity of one of the authors. But between the points of *maximal realizable masked* and unmasked is a plethora of levels.

Studies have shown that status and gender bias are reduced even when author anonymity is rather compromised. Most studies utilized editorial masking, which is successful a little more than half the time. This suggests that one can move along this spectrum, increasing (or reducing) the benefits of DBR while also increasing (or reducing) the cost. The challenge is to figure out where the aggregate of *benefit minus cost* is maximized.

The following proposal attempts to get that balance right. It acknowledges that submissions with less masking have a greater chance of having one or more authors revealed, which might introduce or increase status or gender bias during the review. But also acknowledged is that *uncertainty* as to the authors' identity is often sufficient to realize most or all of the benefits of masking.

3.3.1 *Applicability.* Reviewing of all submissions to *TODS* will be double-blind. All submissions to *TODS* must be masked by the author(s), following the simple instructions given in the proposed author guidelines (Appendix A).

Rationale

Author masking avoids the procedural difficulties of editorial masking, distributes the workload, and, most importantly, significantly increases the efficacy of masking. The literature analysis in the companion paper found that "most of the studies discussed here utilize editorial blinding, which has been shown to be successful about 60% of the time, across many disciplines. Removing text allusions and self-citations would increase success rates to perhaps 75%." [Snodgrass 2006, page 18] We note that the *SIGMOD* conference also utilizes author masking.

Editorial blinding has been shown to be effective in the past, and more aggressive author blinding is even more so. Blank concluded "On the one hand, a substantial fraction—almost half—of the [editorial] blind papers in this experiment could be identified by the referee. This indicates the extent to which no reviewing system can ever be fully anonymous. On the other hand, more than half of the papers in the blind sample *were* completely anonymous. A substantial fraction of submitted papers are not readily identified by reviewers in the field." [Blank 1991, pp. 1051–2] No masking process can be totally successful, yet many of the benefits of DBR accrue from uncertainty in the identity of the author.

This analysis is based on scientific studies that were performed before the advent of the Internet and search engines. In the studies, about half the reviewers were able to guess the identity of at least one of the reviewers. This is partially because

the studies utilized editorial masking. So while author masking does better, the availability of web searches decreases the efficacy, so they balance out somewhat, though there is no good data on the efficacy nor on the impact of status or gender bias in today's more highly-connected environment.

The proposed author guidelines are quite specific on how to ensure anonymity, how and when to use anonymous citations, specifics on papers involving well-known or unique systems, and anonymity in revisions.

The rules for anonymous citations requires such citations both for the authors' own work and that of others, for certain categories of papers (primarily, those that only the author would be aware of). This was done originally in the *SIGMOD* conference author instructions to ensure that reviewers could not infer, from what was anonymous and what was not, the identity of the authors. Say that the paper referenced two papers that it relied heavily on or built upon. One is by author A and the other is an anonymous citation. If that anonymous citation was not masked, but rather was of author B, of a paper to appear, the reviewer could guess that an author of the submitted paper was B, because who else could have built on work that hasn't even appeared? That is the rationale behind requiring anonymous citations even of work by others that is to appear.

As noted before, *TODS* policy requires that submissions invited from conferences utilize some of the same reviewers as the conference submission. Those reviewers will of course know the identity of the authors of the conference paper. However, the other reviewers of the submitted paper need not be told exactly who the authors are. Anonymizing the rest of the journal submission also introduces uncertainty in the reviewer's mind. Were authors dropped or added to this submission? Was the extension carried out by the senior author or by a junior graduate student?

An alternative to the mandatory policy for the remaining submissions is *elective DBR*, where the author decides whether to blind their submission. The experience with elective DBR in other journals is that it works well when there is an established tradition in the community for DBR (as in American Psychology Association, about half of whose journals employ mandatory DBR and most of the rest employ elective DBR) and that it is not invoked often in communities without such a tradition, such as physics,[4] perhaps because authors feel that such submissions will be judged more harshly.

That said, there are some papers in computer science and particularly in database systems that are effectively impossible to blind, and it would be unfortunate if such papers were disallowed. Note however that the instructions to the authors do not require complete blinding, which is in any case not possible for any paper. Rather, all that is required is to follow the simple six steps in those instructions. For some papers, these steps will only partially obscure the authorship, which is understood. Uncertainty of authorship is still preferable to complete knowledge.

A previous version of this proposal placed a higher bar on the degree of obfuscation of the submitted manuscript, but also stated three exception categories:

—Submissions invited from conferences,

[4]`http://forms.aps.org/author/ndbm.pdf` The *IEEE Transactions on Knowledge and Data Engineering* is a more relevant example. This journal employs elective DBR, which is minimally used—perhaps 5% [Werner 2006].)

—Submissions that are extensions of papers that have been previously published in conferences, and

—Submissions for which blinding is effectively impossible.

As several noted, such categories have the potential to create a double standard and cause inconsistencies. Also for such papers it is still important to emphasize to the reviewer that the identity of the author should not influence their review. Hence, in this proposal the exceptions were dropped and the standards for masking reduced.

3.3.2 *Submitting a Manuscript. TODS*, as with most ACM journals and transactions, uses ScholarOne's web-based ManuscriptCentral manuscript tracking system to automate submission and reviewing.

The manuscript will be submitted normally. The EiC will examine the manuscript to ensure that it has been properly blinded.

The cover letter and other correspondence from the author(s) will be marked in ManuscriptCentral (MC) by the EiC as not to be revealed to the reviewers. The handling AE will have full access to these materials.

As mentioned in the author guidelines (Appendix A.2), the cover letter should contain the full details of each anonymous citation as well as a list of people who constitute a conflict of interest.

Rationale

It is important that the EiC and AE have available the details of the anonymous citations for the decisions described below.

A list of conflicts of interest in the cover letter will help avoid cases where a reviewer gets well into a review before realizing that (s)he knows an author and has a conflict with that author. The names should be categorized; there have been cases where authors put under CoI reviewers perceived to have a high reviewing standard.

The *SIGMOD* conference has required, for several years, that conflicts of interest be stated, in a restricted fashion: one must indicate when submitting a paper which program committee member(s) have a conflict of interest, so that the paper will not be assigned to them. The *SIGPLAN* conferences have a similar requirement.

The definition of conflict of interest proposed in the author guidelines is identical to that from NSF [NSF 2004], which is well-thought-out.[5]

Having the author prepare a list of conflicts of interest may take an hour or so. (For those authors who have submitted an NSF grant, satisfying this requirement will be easy.) Once such a list is created, it is not hard to maintain.

Some of the advantages of this list apply to SBR as well. Some other journals have a similar CoI identification requirement; there doesn't seem to be a strong

[5]Jennifer Widom introduced a specific definition of conflict of interest for the *SIGMOD'05* conference; that definition has continued in subsequent conferences. The *SIGMOD* definition differs from that proposed here in that it considers collaborators for only 24 months to have a conflict, whereas the NSF guidelines state 48 months. The other difference is that the *SIGMOD* conference definition also includes the condition "The PC member has been a co-worker in the same department or lab within the past two years." This was not included in this *TODS* proposal to avoid having the author list all members in the department or lab. In most cases the AE will be able to determine this directly from the affiliation of the possible reviewer.

correlation with single/double-blind reviewing and CoI identification.

3.3.3 *Reviewing.* The identity of the authors will be visible to the EiC and Associate Editor; this meta-data will not be revealed by MC to the reviewers. Templates of email to be sent to reviewers will be modified to eliminate identifying information.

Reviewers should direct questions about a paper referenced by an anonymous citation within the manuscript (say, arising from a consideration of the *TODS* novelty requirement) to the AE, who would then provide the cited paper (or equivalently, a full citation) to the reviewer. Instructions to reviewers are given in Appendix B.

In cases where an AE has an intransigent referee who does not respond, the AE sometimes just does the review him/herself, as permitted in the AE manual [Snodgrass 2005]. While this is strictly speaking a violation of DBR, the AE will continue to have this option available.

Rationale

By removing all author names and affiliations from email templates, and by simply removing author names and affiliations from the web pages seen by reviewers, these email templates and web pages can be used for double-blind reviewing.

Anonymous citations are handled similarly to the way they were handled in *SIGMOD'06*, which worked well for the few requests that were made. As each AE handles only a few *TODS* submissions each year, responding to reviewer requests should not constitute a burden.

The instructions to the reviewers does not preclude them from doing web searches or other searches and does not disqualify them if they do discover one of the author's identity. Attempts to prevent such discovery in every single case are bound to fail. Rather, reviewers are simply relied upon not to go to unusual lengths to try to discover the identity of the author. Again, as the culture of the database community becomes more familiar and comfortable with the concept and process of DBR, things should settle out.

Note that the cases where reviews are provided a full citation does not necessarily reveal the full authorship of the submitted paper. The omitted author list on the title page still signals to the reviewer not to utilize that information in their assessment.

3.3.4 *Assessing the Disclosure and Novelty Requirements. TODS* has specific guidelines[6] for disclosing related work by the author(s) of the submitted paper and for what represents adequate contribution over existing published work.

The above procedures allow the disclosure and novelty requirements to be checked. It is completely up to the AE to decide which reviewers should be told about an anonymous citation.

Rationale

TODS policy already requires a *prior publication policy prescreen* by the AE. The cover letter and the submission itself will provide sufficient information for the AE to perform the prescreen.

This is especially important for papers that are extended from conference papers. A recent study, which looked at the papers themselves, found 14 (non-invited)

[6] `http://www.acm.org/tods/Authors.html#PriorPublicationPolicy`

papers extended from conference papers published in *TODS* over calendar 2003–2005 [Montesi 2006]. There were also 22 papers invited from conferences (mainly *SIGMOD* and *PODS*) that were published during that three-year period. This works out to 56% of published papers. However, as invited submissions have a much higher success rate, it is estimated that slightly less than half of the submissions to *TODS* are extended from conference papers.

The disclosure requirements (30%) for papers that are extensions of conference papers have in the past been generally checked by reviewers. That will continue under DBR, in two phases. First, the AE will check the paper to ensure that statements about contributions beyond prior published work (e.g., "We added a section on algorithms", "we now present a detailed proof of Theorem 4"), consulting with the referenced papers, whose citations are in the cover letter. The second phase is the actual determination of degree of additional contribution by the reviewers. This policy proposes that should a reviewer have questions, the AE is permitted to reveal the citation to the reviewer (thus partially unblinding the paper), so that the reviewer can fully check that the novelty requirement has been fulfilled.

Note however that this has the effect of devolving perhaps a good percentage of reviews of such papers to SBR. However, the author identity will still be omitted from the cover page of the submitted paper, reminding the reviewer that the identity of the author should not influence their review. The justification is that comprehensiveness of the review is judged to be more important than blinding efficacy.

3.3.5 *On Revision.* The author guidelines includes a section on anonymity in revisions (Appendix A.5).

Rationale

Until the paper is accepted, author identities must remain masked.

3.3.6 *Acceptance.* As the AE is considering whether to accept the manuscript for publication, they will consider the novelty requirements, which for double-blinded reviews necessarily involve anonymous citations, which as noted above must be fully documented in the cover letter.

If and when the manuscript is accepted for publication, the author will prepare a non-anonymized version. The AE will at that time make a final check to ensure that the *TODS* disclosure requirements are met.

Rationale

The specifics of meeting the disclosure requirements can be checked with the final version of the paper. If there are any concerns, they can be easily dealt with at that point.

3.3.7 *Transition.* It is important to educate the community as to the benefits and costs of DBR. It is also important to publicize and promulgate policy changes widely, *before* those changes go into effect. Finally, it is critical to ensure that ManuscriptCentral is handling things correctly before adopting such policy changes.

Rationale

The transition to DBR is a significant one for a journal, and so must be done carefully. Such changes require time for the community to understand and assimilate. The experience from the *SIGMOD* conference is that it can take years for the

community to come to embrace DBR.

3.4 Cost-Benefit Analysis

We revisit here the benefits and costs listed in Sections 3.1 and 3.2, respectively.

DBR avoids a bias towards top-ranked authors and institutions. The experimental evidence is mixed concerning status bias. There is no clear advantage to top-ranked authors and to authors from top-ranked institutions of SBR, nor a clear disadvantage of DBR.

DBR avoids a bias towards prolific authors. The contradictory results of prior studies render it impossible to say anything definitive about the impact of blinding on prolific authors.

DBR encourages higher-quality reviews. It has not been shown convincingly that either SBR or DBR can increase the quality of reviews of a submitted manuscript.

DBR is more fair to women authors. As all submissions are under DBR, gender bias should be less of a concern.

DBR is more fair to non-top-ranked authors and institutions. Similarly, status bias should be less of a concern.

There is a perception by some that DBR is more fair. Requiring DBR of all submissions sends a strong message to the community that *TODS* is making a sincere attempt to fully meet ACM's promise of an appropriate and timely decision based on proper review by well-qualified and impartial reviewers.

We now turn to an analysis of the costs of DBR of implementing the policy just specified.

Quality of reviews is reduced. It has not been shown definitively that DBR reduces the quality of reviews.

Intentional bias is reduced. While there is substantial evidence of referee bias in SBR, there is little evidence that referees were intentionally biased in carrying out their duties as reviewers. Rather, the bias that was observed seems to be implicit and unintended.

It is difficult to mask submitted manuscripts. With author masking, the source of the manuscript is available and the blinding is done by the person most familiar with that manuscript, that is, the author. Detailed guidelines and examples are provided in the author instructions. It is anticipated that blinding a manuscript should take no more than an hour of work.

We acknowledge that, in the presence of web search engines, submitted manuscripts will be successfully masked less often than before such search engines were available. The proposed approach does not require nor depend on complete masking.

DBR increases the probability of non-novel papers getting accepted. In the reviewing procedure described here, reviewers are free to search for related work. The responsibility for determining the contribution of the work remains, whether the reviewing is single- or double-blind.

It is difficult to fully mask systems papers. The required anonymization steps are easy to apply to all papers. Full masking is indeed difficult in some cases, and so

is not required. All that is required is to follow the simple six steps listed in the author instructions.

DBR jeopardizes "power projects". The submitted manuscript is of course free to mention the project upon which it is based. What is not permitted is to state or imply that an author of the submitted manuscript is one of the developers of that original project.

As far as the reviewer knows, it is possible that another group has acquired the software and has extended it, as exemplified in the author instructions. Perhaps the author is a new graduate student in the research group, or a student who has graduated and is at a different institution, or is a friend of someone in the research group and was thus able to acquire the system through that means.

All that is necessary is to introduce uncertainty as to the identity of the author.

Of course, the author will be clearly identified when the paper is published, thereby extending and strengthening the "brand" of these power projects.

Forward-looking systems work, in particular, suffers from anonymization. It is recognized that some submissions reporting work on forward-looking systems may not be truly anonymous.

Recall that the objective is to introduce uncertainty, not preclude all possible ways of guessing the author.

In summary, it seems that the proposed process does not in any way put systems papers, papers on power projects, or forward-looking papers at a disadvantage. They are allowed to compete with other papers on an equal footing; anonymization that obscures the scientific contributions of such papers (or any other papers, for that matter) is not required.

Anonymization is a distraction to students learning to write clearly. The author guidelines give simple steps for converting a clearly written paper into one for which the authorship is not revealed. These six steps can be performed in the last hour before a manuscript is submitted to *TODS*.

This requirement provides a useful pedagogical opportunity to explain to students the vagaries of bias, the subtleties of the peer review process, and finally, the need for the *paper itself*, and not extraneous aspects such as the identity or institution of the author(s), to convey the correctness of the arguments.

Some journal policies preclude DBR. The proposed policy shows that DBR can be executed while not jettisoning any existing editorial policies for *TODS*.

DBR precludes expanding a conference paper into a journal submission. The proposed policy fully supports submissions expanded from published conference papers.

DBR makes it difficult for conflicts of interest to be detected. The proposed process makes conflicts of interest explicit in the cover letter. There is an initial cost to the author of preparing such a list. But this cover letter helps to reduce the incremental load for the AEs, as they would not have to avoid such conflicts only through their knowledge of the network of relationships in the research community. (This is an advantage regardless of the SBR-DBR distinction.)

DBR prevents the dissemination of research results before submission. The proposed policy dictates no constraints on such dissemination. The only requirement is that anonymous citations be disclosed to the EiC and AE in the cover letter (not to be shown to the reviewers).

DBR gets papers rejected for failing to sufficiently cite the author's own work. TODS has specific novelty and disclosure requirements that build on the ACM self-plagiarism policy.[7] Not following these guidelines could lead to rejection of the submitted manuscript.

The anonymization guidelines in the author instructions show how the author's own work may be cited. Doing so allows the author to simultaneously satisfy the DBR, novelty, and disclosure requirements. The proposal mechanism states when each of these requirements is evaluated and ensures that adequate information is available at the appropriate time.

DBR increases the possibility of plagiarism. In the proposed mechanism, reviewers are still responsible for detecting plagiarism. As the community becomes more fluent with DBR, reviewers will not assume that they know who all the authors are.

DBR raises the possibility of self-plagiarism. The above analysis and that in Section 3.3.4 argue that there are still effective processes in place for detecting self-plagiarism.

DBR places extra burdens on reviewers. The six steps to blind a paper shouldn't affect the readability of the paper.

Novelty is somewhat harder to check for a blinded submission. Should the reviewer wish to check a reference that has been anonymized, they can ask questions of the AE, who is permitted if needed to reveal the citation to that reviewer.

DBR decreases enthusiasm of reviewers. There is little evidence from the literature that a policy of DBR reduces incentives to do reviewing, reduces respect for reviewers, or makes it more difficult for a journal to find reviewers. There may be an initial transient effect when DBR is first adopted, but that should decrease as the community becomes comfortable with such a policy.

DBR reduces timeliness. A previously-proposed protocol for checking novelty, requiring several interactions between a reviewer and an AE, has now been simplified: if needed, the reviewer just asks the AE for the full citation of an anonymous citation.

DBR places extra burdens on the EiC and AEs. In the proposed process, the EiC has to configure MC so that it doesn't reveal the authors to the reviewers in its dynamic web pages or email templates. The EiC also must convey the new procedures to the community and answer the inevitable questions. Finally, the EiC must examine each submission with a quick scan to ensure that it has been properly blinded.

DBR also imposes additional work on AEs: they have to respond to questions from reviewers regarding the material referenced by anonymous citations. On the other hand, the list of conflicts provided by the authors makes assigning reviewers slightly easier on AEs.

DBR places extra burdens on the authors. The authors are required to spend perhaps an hour following the steps in the author instructions to anonymize their paper, as well as perhaps an hour preparing the cover letter with the list of conflicts of interest, if that list is not already available.

[7] `http://www.acm.org/pubs/sim_submissions.html`

SBR versus DBR should not be a local decision. As the previous section noted, the prevalence of DBR has increased dramatically over the last fifteen years, to the point where most scientific journals now employ double-blind reviewing. Closer to home, an informal vote at the SIGMOD 2006 business meeting in Chicago indicated wide-spread support for DBR. A survey of database researchers [Sirbu and Vasilescu 2006] listed in the final recommendations "To maintain the perception of fairness (even though it might not have a large impact on the acceptance rates), double-blind review could be useful." [page 58].

Perhaps consistency is a benefit of DBR rather than a cost.

Looking at this another way, at the present time a database researcher has absolutely no choice, as none of the journals currently employ mandatory DBR. Having *TODS* adopt DBR provides choice to authors.

We have analyzed over twenty possible costs to DBR. Many of these costs are completely eliminated through provisions of the proposal; examples include jeopardizing power projects and forward-looking systems, journal policies inconsistent with DBR, precluding expanding conference papers into *TODS* submissions, and preventing dissemination of results before (or after) submission. It seems that the primary costs are additional burdens on the EiC, AEs, and authors. The burdens on the EiC and AEs are justified by the desire of *TODS* to provide a reviewing process that is as fair as possible. The burden on the authors, one or two hours per submission, is the price for ensuring that gender bias, status bias, or other bias is less likely to be a factor in the review of their submission.

We see that the administrative cost can be lowered through a combination of author anonymization, system support (via MC), editor judgment in specific cases (by the EiC for initial submission, and by the handling AE for anonymous citations), and carefully-worded guidance that provides a specific set of steps that can easily be executed by authors and reviewers.

4. DISCUSSION

(While the literature survey and the cost-benefit analysis above attempt to be objective, the present section presents the author's personal view.)

As the noted statistician Lynne Billard, who has written extensively on this topic, has remarked, "The issue of double-blind refereeing today is one fraught with emotional overtones both rational and irrational, often subconsciously culturally based, and so is difficult for many of us to resolve equitably no matter how well intentioned" [Billard 1993, page 320].

4.1 Why does this matter to *TODS*?

The ACM Publications Board approved several years ago a broad rights and responsibilities policy [ACM 2001; Snodgrass 2002]. Included was the promise that "The aim of the review process is to make an appropriate and timely decision on whether a submission should be published. Such decisions are based on proper review by well-qualified and impartial reviewers. ... Thus authors can expect ACM to use impartial reviewers and to issue timely review and clear feedback."

Each year the *TODS*' associate editors render editorial decisions for about one hundred submitted manuscripts. The editors and reviewers are for the most part

diligent and careful in carrying out the editorial process. That said, the scientific studies reviewed here and in the companion paper suggest that for a small percentage of those submissions, the wrong editorial decision was made.

It should be emphasized that *which* papers were erroneously rejected (or accepted) is not known. Referees never write "this submission can't be good: look where it came from." They don't even think that consciously.

The impetus for these incorrect editorial decisions lies in an inherent bias in the reviewing process that *TODS* uses [Kardes et al. 2005; Wilson & Brekke 1994]. This process provides irrelevant, distracting information to reviewers that causes these reviewers, through no intention or conscious effort, to see the paper in a more negative light.

No study has been done of the fairness of any ACM journal or transaction. But very careful studies have been done of similar journals. These studies have found systematic bias in unmasked reviewing. The probability of such wrong decisions is evident only because reviewer bias is so well-documented in the literature.

That this has occurred is not the fault of the *TODS* volunteer editors. It is not the fault of the hundreds of reviewers, who are professional, caring, objective scholars. It is the fault of a *process* that is patently unfair, due to the various ways that bias psychologically affects judgment.

An especially productive way to avoid mental contamination of the judgment process is to control cones exposure to biasing information, termed *exposure control* [Wilson & Brekke 1994]. In terms of scholarly review, exposure control is accomplished by blinding the reviewer to the identity and affiliation of the author.

The literature also shows that ACM is an outlier in the scientific community, in that it uses SBR for all of its journals. The majority of scientific journals now use DBR. And it is clear that the more a scientific discipline studies bias in general, the more it utilizes DBR.

4.2 Does bias exist in computer science?

Bias in judgment is inevitable in complex tasks with many dimensions [Wilson & Brekke 1994]. Journal reviewing is a prime example of such a task. Major studies, described in Section 2.2, found evidence of bias in economics and medical research. It is true that some submissions to *TODS* are highly mathematical and thus might be less prone to reviewer bias. However, the review of even highly mathematical papers requires subjective judgment as to the contribution of the paper, a judgment in which reviewer bias is certainly possible.

In the absence of a compelling countervailing mechanism, a reasonable conclusion is that the *TODS* reviewing process, like that of the journals for which those studies were done, is systemically unfair, despite the promises of the afore-mentioned ACM policy.

4.3 What is the appropriate decision procedure?

The ACM rights and responsibilities policy puts the onus on ACM and on the EiC to ensure that the reviewing process is fair.

Thus the issue is not, how to make the reviewing fair when it is easy to administer. *TODS* goes to great lengths, at great cost, to ensure fair reviewing. Reviewer identity is held in confidence. The review form lists several aspects to consider,

to guide the reviewer in considering appropriate aspects. Reviews are much more than just recommendations (accept/reject): reviewers are expected to document in detail their rationale for their recommendation. Several reviewers are used for each paper, to not be overly reliant on just one or two views. These elaborate procedures require months per submission. An editorial decision could be made much more easily, in just a few days or weeks, if fairness wasn't viewed as critical.

The appropriate decision procedure is to first decide, is SBR or DBR more fair? (This is assuming that DBR can be done at all, but many journals have shown by example that DBR is possible with journals.) And only then, if DBR is judged to be more fair, to determine how the process can be made as smooth as possible and with as little cost as possible, while retaining fairness and balancing remaining costs based on stated principles.

Some of the costs can be reduced unilaterally, as shown in Section 3.3. For example, journal policies that preclude DBR can be adjusted to accommodate DBR and conflicts of interest can be identified even with DBR.

Some of the costs are discipline-specific. For disciplines that do not involve "power projects," such as the more mathematically-oriented disciplines, a provision for ambiguating statements on well-known or unique systems may not be relevant. For disciplines in which there is not a tradition of extending conference papers into journal submissions, a procedure for revealing anonymous citations may not be needed. These are "knobs" that a policy designer can adjust to fine-tune the policy.

But many of the costs, if reduced, increase other costs, due to complex interactions. The approach taken in Section 3.3 was to apply a principled tradeoff between such costs and blinding efficacy. These are the other "knobs" that are available to policy designers.

Three principles are used here. The first is that authors should not be required to go to great lengths to blind their submissions. The second is that comprehensiveness of the review trumps blinding efficacy. The final principle used here is that AEs retain flexibility and authority in managing the reviewing process. Hence, the procedure dictates six quick steps for blinding a submission, when a more aggressive obfuscation of the paper would yield higher blinding efficacy. And the procedure explicitly allows the AE to reveal author identity information, such as a full citation that had been previously anonymized, if needed by a reviewer to evaluate the submission.

Applying these principles encouraged a systematic design of the tradeoff between costs and blinding efficacy.

4.4 Decision

In considering this decision, it is my responsibility as Editor-in-Chief to understand the scientific data and to gather input from all constituencies. I have read most of the papers concerning blind review and have talked to perhaps one hundred people.

The data shows that status bias is against those in the gray area: neither at the top (prominent, prolific, or from top schools: the *scholarly elite*) nor at the bottom (those who can't write a good paper). It is those in the gray area who are hurt by the current system.

I involved the *TODS* editorial board because it is the AEs who must execute the

submission handling process. I worked with the editorial board for two months to make the process as smooth as possible. The AEs are the experts on the process and have provided many excellent suggestions that have resulted in the clean, workable design presented in Section 3.3.

Concerning the scholarly elite, I consider the *TODS* editorial board to be a representative sample. The consensus among the *TODS* editorial board is that SBR is less fair than DBR. When asked for their recommendation, roughly a third recommended SBR, roughly a third recommended DBR, and roughly a third were comfortable with either SBR or DBR.

But the primary constituency consists of qualified potential authors, those who write reasonable papers but who suffer from status bias from a flawed process. Note that none of these people is on the editorial board.

An informal vote taken at the SIGMOD business meeting in Chicago in June 2006, in a room filled with hundreds of qualified potential authors, was decisive: over 90% were in favor of DBR. That vote is mirrored in the individual discussions I have had. Most are astounded that there is even a question.

Given this information, my decision was a natural one. The many qualified potential authors, for whom SBR may be less fair, overwhelmingly favor DBR. The scholarly elite, who are small in number and who might in some cases unfairly benefit from the current process, are mixed. And it has been shown to be possible to design an editorial process that minimizes the costs while still emphasizing the primacy of the merits of the submission.

My conclusion is simple: the *TODS* reviewing process should be fair to everyone.

Given that the community has spoken with such a clear voice and that there are so few instances of journals in which mandatory DBR has been subsequently replaced with SBR, I believe that this conclusion will be embraced by the vast majority of the community, as authors and reviewers and AEs gain experience with the reviewing process.

5. TODS POLICY

The following is now *TODS* policy.

—*TODS* reaffirms the general ACM policy that "the quality of a refereed publication rests primarily on the impartial judgment of their volunteer reviewers."

—*TODS* will continue to strive to ensure fairness in reviewing, even if that involves more work for the *TODS* editorial board.

—Scientific studies have demonstrated opportunities for bias inherent in single-blind reviewing.

—It is *TODS* policy that every submission should be judged on its own merits. The identity and affiliation of the authors should not influence, either positively or negatively, the evaluation of submissions to *TODS*.

—For this reason, *TODS* will utilize double-blind reviewing.

—*TODS* will continue to strive to make the submission process for authors as simple as possible.

—*TODS* will continue to strive to effect a comprehensive review of each submission.

This policy is not dependent on absolute or even relative blinding efficacy. The central and unambiguous message is that every submission should be judged solely on its own merits. This message applies even when reviewers know exactly who the authors are. The other important message is that *TODS* so values fairness that it is willing to undertake additional effort by AEs to make the process more fair.

Acknowledgments

The instructions to authors and instructions to reviewers given in the appendix are based on instructions given to *SIGMOD'06* authors. These instructions have benefited from editing by the program chairs of *SIGMOD'01* through *SIGMOD'06*: Timos Sellis, Michael Franklin, Alon Halevy, Gerhard Weikum, Jennifer Widom, and Surajit Chaudhuri.

The author thanks the many *TODS* associate editors and others who have corresponded privately with their ideas and concerns about SBR and DBR, including but not limited to Kurt Akeley, Alan Berenbaum, Judy Brown, Merrie Brucks, Boots Cassel, Surajit Chaudhuri, Jan Chomicki, Douglas D. Dankel II, Peter Denning, David DeWitt, Susan Dumais, Curtis Dyreson, Kemal Ebcioglu, Carla Ellis, Mary Fernandez, Michael Franklin, Richard P. Gabriel, Luis Gravano, Ralf Güting, Alon Halevy, Joe Hellerstein, Christian S. Jensen, Susan Jones, Won Kim, Phokion Kolitis, Hank Korth, Donald Kossmann, David Leake, Steven Levitan, Alan Mackworth, Kathryn McKinley, Elli Mylonas, Fred Niederman, Beng Chin Ooi, Z. Meral Özsoyoğlu, Tamer Özsu, Craig Partridge, Ray Perrault, Michael Pristley, Raghu Ramakrishnan, Stephanie Rosenbaum, Kenneth Ross, Betty Salzberg, Pierangela Samarati, Sunita Sarawagi, Stuart Selber, David E. Siegel, Dan Suciu, S. Sudarshan, Geoff Sutcliffe, Marilyn Tremaine, Jan Van den Bussche, Mary Vernon, Jennifer Widom, Nina Wishbow, and Xindong Wu. Their comments have greatly improved this analysis.

REFERENCES

ACM Publications Board, Rights and Responsibilities in ACM Publishing, approved June 27, 2001. `http://www.acm.org/pubs/rights.html`, viewed December 5, 2006.

N. Altman, D. Banks, P. Chen, D. Duffy, J. Hardwick, C. Léger, A. Owen, and T. Stukel, Meeting the Needs of New Statistical Researchers, *Statistical Science* 6(2):163–174, May 1991.

N. Altman, J. F. Angers, D. Banks, D. Duffy, J. Hardwick, C. Léger, M. Martin, D. Nolan (Chair), A. Owen, D. Politis, K. Roeder, T. N. Stukle and Z. Ying, Rejoinder, *Statistical Science* 7(2):265–266, May 1992.

L. Billard, Comment, *Statistical Science* 8(3):320–322, August 1993.

R. C. Calfee and R. R. Valencia, APA Guide to Preparing Manuscripts for Journal Publication, `http://www.apa.org/journals/authors/guide.html`, viewed June 21, 2006.

R. M. Blank, The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from The American Economic Review, *American Economic Review* 81(5):1041–1067, December 1991.

G. Borts, Report of the Managing Editor, *American Economic Review* 64:476–482, May 1974.

S. J. Ceci and D. P. Peters, How blind is blind review? *American Psychologist* 39:1491–1494, 1984.

D. Cox, L. Gleser, M. Perlman, N. Reid, and K. Roeder, Report of the Ad Hoc Committee of Double-Blind Refereeing, *Statistical Science* 8(3):310–317, August 1993.

M. Dalton, Refereeing of Scholarly Works for Primary Publishing, in *Annual Review of Information Science and Technology (ARIST)*, Volume 30, M. E. Williams (ed.), American Society of Information Science, 1995.

J. M. Garfunkel, M. H. Ulshen, H. J. Hamrick, and E. E. Lawon, Effect of Institutional Prestige on Reviewers' Recommendations and Editorial Decisions, *Journal of the American Medical Association* 272(2):137–138, July 13, 1994.

C. Genest, Comment, *Statistical Science* 8(3):323–327, August, 1993.

M. D. Gordon, The role of referees in scientific communication, in *The Psychology of Written Communication: Selected Readings*, J. Hartley (ed.), London, England: Kogan Page, pp. 263–275, 1980.

A. G. Greewald and M. R. Banaji, Implicit social cognition: Attitudes, self-esteem, and stereotypes, *Psychological Review*, 102(1):4–27, January 1995.

D. F. Horrobin, Peer review: A philosophically faulty concept which is proving disastrous to science, *Behavioral and Brain Sciences* 5(2):217–218, 1982.

F. R. Kardes, A. V. Muthukrishnan, and V. Pashkevich, On the conditions under which experience and motivation accentuate bias in intuitive judgment, in *The Routines of Decision Making*, T. Betsch and S. Haberstroh (ed.), Lawrence Erlbaum Associates, pp. 139–156, 2005.

A. M. Link, US and Non-US Submission: An Analysis of Review Bias, *Journal of the American Medical Association* 280(3):246–247, July 15, 1998.

M. J. Mahoney, A. E. Kazdin, and M. Kenigsberg, Getting Published, *Cognitive Therapy and Research* 2(1):69–70, 1978.

M. Montesi, From Conference to Journal Publication. How Conference Papers in Software Engineering are Extended for Publication in Journals, in preparation, 2006.

National Science Foundation, Conflict-of-Interests and Confidentiality Statement for NSF Panelists, NSF Form 1230P, February 2004.

A. Rosenblatt and S. A. Kirk, Recognition of Authors in Blind Review of Manuscripts, *Journal of Social Service Research* 3(4):383–394, Summer 1980.

B. Shiv, Z. Carmon, and D. Ariely, Placebo Effects of Marketing Actions: Consumers May Get What They Pay For, *Journal of Marketing Research* 42(4):383–393, November 2005.

A. Sirbu and A.-T. Vasilescu, The Health of Database Research, Project Report, Worcester Polytechnic Institute, May 2006.

R. T. Snodgrass, Rights and Responsibilities in ACM Publishing, *CACM* 45(2): 97–101, February 2002.

R. T. Snodgrass, ACM TODS *Associate Editor Manual*, September 2005.

R. T. Snodgrass, "Single- Versus Double-Blind Reviewing: An Analysis of the Literature," *ACM SIGMOD Record*, 35(3):8–21, September 2006.

S. van Rooyen, F. Godlee, S. Evans, R. Smith, and N. Black, Effect of Blinding and Unmasking on the Quality of Peer Review: A Randomized Trial, *Journal of General Internal Medicine* 14(10):622–624, October 1999.

S. Werner, personal communication, December 4, 2006.

T. D. Wilson and N. Brekke, Mental contamination and mental correction: Unwanted influences on judgments and evaluations, *Psychological Bulletin* 116(1):117–142, July 1994.

H. Zuckerman and R. K. Merton, Patterns of Evaluation in Science: Institutionalization, Structure and Functions of the Referee System, *Minerva* 9(1):66–100, January 1971.

A. INSTRUCTIONS TO AUTHORS

TODS strives to ensure fairness in reviewing. It is *TODS* policy that every submission should be judged solely on its own merits. The identity and affiliation of the authors should not influence, either positively or negatively, the evaluation of submissions to *TODS*. For this reason, all submissions to *TODS* will undergo

double-blind review, in which authors and reviewers are unaware of each other's identities.

A.1 Ensuring Anonymity

To ensure anonymity of authorship, authors must blind their manuscript by performing the following simple alterations.

—Authors' names and affiliations must not appear on the title page or elsewhere in the paper.

—Funding sources(s) must not be acknowledged on the title page or elsewhere in the paper.

—All personal acknowledgments should be omitted. Research group members or other colleagues or collaborators must not be acknowledged anywhere in the paper. There should also be no acknowledgment section in the paper.

—Source file naming must also be done with care. It is strongly suggested that the submitted file be named as `todssubmission.pdf`. Also, if your name is Jane Smith and you submit a PDF file generated from a .dvi file called `Jane-Smith.dvi`, one can infer your authorship by looking into the PDF file.

Despite the anonymity requirements, you should still include relevant prior published work of your own in the references—omitting them could potentially reveal your identity by negation. *Prior published work* is defined as any research paper that has been published and made available prior to submission to *TODS* (a) in the online or printed proceedings of a refereed conference or refereed workshop, (b) as a longer poster papers (4 pages or more) in such a conference or workshop, or (c) in an online or printed issue of a journal. However, you should not cite a published demonstration paper of your own even if it is in prior published work as defined above.

You must use care in referring to your prior published work. For example, if you are Jane Smith, the following text reveals the authorship of the submitted paper:

> In our previous work [Doe 1997, Smith 1998], we presented two algorithms for .... We build on that work by ...
>
> Bibliography
>
> John Doe and Jane Smith, A Simple Algorithm for ..., *Proceedings of ACM SIGMOD International Conference on Management of Data*, pp. 1–10, 1997.
>
> Jane Smith, A More Complicated Algorithm for.., *Proceedings of ACM SIGMOD International Conference on Management of Data*, pp. 34–44, 1998.

The solution is to reference your prior published work in the third person (just as you would any other piece of work that is related to the submitted paper). This allows you to set the context for the submitted paper, while at the same time preserving anonymity.

> In previous work [Doe 1997, Smith 1998], algorithms were presented for ... We build on that work by ...
>
> Bibliography

JOHN DOE AND JANE SMITH, A Simple Algorithm for ..., *Proceedings of ACM SIGMOD International Conference on Management of Data*, pp. 1–10, 1997.

JANE SMITH, A More Complicated Algorithm for.., *Proceedings of ACM SIGMOD International Conference on Management of Data*, pp. 34–44, 1998.

A.2 Anonymous Citations

Referring to any work (including but not limited to your own work) that has been submitted elsewhere for review (and hence is as yet unpublished), or that has been accepted for publication at a referred conference, referred workshop, or a journal for which proceedings (printed or online) will be made available after the submission of your *TODS* manuscript, requires a different protocol to ensure consistency with double-blind reviewing. In the body of your *TODS* submission, you may refer to such work in the third person as follows.

> The authors have also developed closely related techniques for query optimization [Anonymous], but...

In the above example, the reference in the bibliography would then read:

> ANONYMOUS, details omitted due to double-blind reviewing.

Note that you should mention neither the authors nor the title or venue of publication while describing *anonymous citations* like that above. Be sure to place all anonymous citations after the list of your regular citations. Of course, this does not mean you are not responsible to provide the details of such anonymous citations. You should disclose (for use by the Editors only) full details of each such anonymous citation in your cover letter, which will not be made available to the reviewer. Furthermore, you may be asked to submit a copy of one of these papers corresponding to such citations. An editor will contact you during the review period if this becomes necessary.

Technical reports (or URLs for down-loadable versions) of your own work should not be referenced. Self-references should also be limited to only papers that are relevant and essential for the reviewing of the submitted paper.

A.3 Conflicts of Interest

As conflicts of interest between the author and reviewer are harder to detect in double-blind review, authors are required to submit a list of all such conflicts in their cover letter. To reduce the effort of preparing such a list, *TODS* adopts the definition of conflict of interest utilized by the US National Science Foundation.

Specifically, you should list all people who are potential reviewers and who have a relationship with one of the authors involving the following.

—Known family relationship as spouse, child, sibling, or parent.

—Business or professional partnership.

—Past or present association as thesis advisor or thesis student.

—Collaboration on a project or on a book, article, report, or paper within the last 48 months.

—Co-editing of a journal, compendium, or conference proceedings within the last 24 months.

—Other relationship, such as close personal friendship, that you think might tend to affect your judgment or be seen as doing so by a reasonable person familiar with the relationship.

Please list these people under the appropriate categories.

A.4 Well-Known or Unique Systems

Papers on well-known or unique systems may be more difficult to blind. All that is required is that blatant statements that could identify one of the authors be ambiguated.

A common approach is to state that the developers of the system (which might well have been the authors of the submitted paper) made the system available to the authors of the submitted paper (which is strictly a true statement). For software, the developers might have shared the source code with the authors; for a hardware system, the developers might have made the system available for use by the authors. It might even be useful to explicitly mention that the developers provided guidance on the internals of the system.

There is no expectation that manuscripts be extensively rewritten to attempt to mask the identity of the authors.

A.5 Anonymity in Revisions

All comments to the reviewers and all manuscript revisions except for a possible final revision after the paper has been accepted (should the paper be accepted) should also follow the guidelines above.

A.6 Summary

It is the responsibility of authors to do their very best to follow the specific steps listed here to preserve anonymity. Papers that do not follow the guidelines here are subject to immediate rejection.

The steps can be summarized quickly.

(1) Anonymize the title page.
(2) Remove mention of funding sources and personal acknowledgments.
(3) Anonymize references found in running prose that cite your papers.
(4) Anonymize citations of submitted work in the bibliography.
(5) Ambiguate statements on well-known or unique systems that identify an author.
(6) Name your files with care.

Authors need only take these six steps to adequately blind their papers.

Common sense can go a long way toward preserving anonymity without diminishing the quality or impact of a paper. The goal is to preserve anonymity to a reasonable degree while still allowing the reader to fully grasp the context (related past work, including your own) of the submitted paper and while making it relatively easy for the author to follow the required steps. If you desire specific guidance, please contact the Editor-in-Chief.

B. INSTRUCTIONS TO REVIEWERS

It is *TODS* policy to use *double-blind reviewing* for all papers, in which authors and reviewers are unaware of each other's identities. It is *TODS* policy that every submission should be judged solely on its own merits. The identity and affiliation of the authors should not influence, either positively or negatively, the evaluation of submissions to *TODS*. In particular, you are asked not to go to unusual lengths to try to discover the identity of the author. On the other hand, you should still be diligent in determining the contribution of the submission over previously-published work by the author or by others.

Authors have been given guidelines on how to approach double-blind reviewing. Please review the submission page to familiarize yourself with these guidelines. If authors have not made reasonable effort to conform to double-blind reviewing, the paper could be rejected. "Reasonable effort" means the author followed the guidelines in the submission page. Your ability to guess/discover an author's identity based on your past experience, extensive web search, non-obvious "clues" in the submitted paper, and other exploratory means are not necessarily grounds for rejection.

Should you need access to the material referenced by anonymous citations (those stating "details omitted due to double-blind reviewing"), say to judge the novelty requirement, simply notify the Associate Editor handling this paper and give your reason and the full citation will be revealed to you. If you have any other questions about double-blind reviewing, please also contact the Associate Editor.