Double blind is good but open would be better: Perceptions of peer review in the SE community

Lutz Prechelt
Freie Universität Berlin
Berlin, Germany
prechelt@inf.fu-berlin.de

Daniel Graziotin
University of Stuttgart
Stuttgart, Germany
Daniel.Graziotin@iste.uni-stuttgart.de

Daniel Mendez
Blekinge Institute of Technology and fortiss GmbH
Karlskrona, Sweden
daniel.mendez@bth.se

1. What we did
Peer review in software engineering is considered, same as for other disciplines, to be a key element of the research process, yet it is often perceived as not to work fully well. To understand the pains and gains in the peer review system, we ran a survey with open and closed questions with the authors and PC members of ICSE 2014/2015/2016. We received 241 responses (29% response rate). 67% of the respondents identified themselves as professors. We analyzed the responses quantitatively and qualitatively (with open coding). All questions were optional. Agreement scales had 10 points, so mild levels of agreement could be expressed but there was no undecided middle point. The resulting article appeared in Information and Software Technology in 2018 [1] and we also disclosed the anonymized data set [2].

2. What we found
1. Purpose: The top four purposes our respondents saw for peer review (over 85% agreement) were making sure results are valid, articles are readable, limitations are addressed, and research is relevant. Lowest (59% agreement) was protecting the venue's reputation.
2. Review quality: The fraction of reviews considered "good" quality was typically only one third (but individual answers varied widely), "reasonable" another third, "unhelpful" about 20%, and "grossly faulty" 10%. As the main reasons for unhelpful or faulty reviews, respondents claimed insufficient time spent (24%) or being unfamiliar with the topic area (22%), all other reasons had 10% support or less. The article sketches heaps of ridiculous anecdotes of type "worst review experience" from authors’ perspectives (with typical patterns) and from reviewers’ perspectives (more varied).
3. Blinding: 63% agreed that reviewers should not know authors ("double-blind"); a stunning 36% agreed that authors should know reviewers ("signed reviews"); 52% agreed that reviews should be published along with articles; as many as 31% agreed that submissions and reviews should be published for rejected articles.
4. Reviewer compensation: 41% agreed that reviewers should receive (quasi-)monetary compensation; 14 compensation types were mentioned, with monetary at 16% of support. In contrast, the idea of "showable proof of good reviewing work" had 71% support with 31 different suggested types, such as handing out a certificate (51% support), but perhaps only for the best reviewers (16%) or with an explicit statement about review quality (11%). Other forms of transparency or a reputation system sum to 27% support.
5. Trends: We investigated linear models using age and/or career stage to identify possible generational trends or seniority trends for various potentially dependent responses in turn, filtering for highly significant (p < 0.02) models. Some findings: Review quality perception appears to be trendless; younger respondents prefer to hide their name from co-reviewers, older ones do not; other blinding attitudes are trendless; younger researchers favor (quasi-)monetary compensation, older ones do not; many other responses are trendless. Overall: We should not expect peer review to change just because an older generation of researchers will retire.

3. So what?
Recent moves to double blind review have been met favorably by SE researchers. Yet, review quality leaves a lot to be desired. What else to do? Various non-SE venues show that open peer review is possible and initiatives such as Review Quality Collector (https://reviewqualitycollector.org) promise improvements even in blinded reviewing regimes. Let us go forward!

4. REFERENCES