Can editors protect peer review from bad reviewers?

Rafael D'Andrea[1,*] **and James P O'Dwyer**[1]

[1]University of Illinois, Department of Plant Biology, Urbana-Champaign IL, 61801, USA
[*]rdandrea@illinois.edu

ABSTRACT

Peer review is the gold standard for scientific communication, but its ability to guarantee the quality of published research remains difficult to verify. Recent modeling studies suggest that peer review is sensitive to reviewer misbehavior, and it has been claimed that referees who sabotage work they perceive as competition may severely undermine the quality of publications. Here we examine which aspects of suboptimal reviewing practices most strongly impact quality, and test different mitigating strategies that editors may employ to counter them. We find that the biggest hazard to the quality of published literature is not selfish rejection of high-quality manuscripts but indifferent acceptance of low-quality ones. Blacklisting bad reviewers and consulting additional reviewers to settle disagreements can reduce but not eliminate the impact. The other editorial strategies we tested do not significantly improve quality, but pairing manuscripts to reviewers unlikely to selfishly reject them and allowing revision of rejected manuscripts minimize rejection of above-average manuscripts. In its current form, peer review offers few incentives for impartial reviewing efforts. Editors can help, but structural changes are more likely to have a stronger impact.

Introduction

Peer review is the main process by which scientists communicate their work, and is widely regarded as a gatekeeper of the quality of published research[1]. Surprisingly, its efficiency remains largely assumed rather than demonstrated[2], and the fairness, reliability, and transparency of the process have been repeatedly questioned[3–5]. Given the scarcity of empirical data, modeling approaches are increasingly used to test peer review[6–10], with particular attention devoted to the impact of referee behavior on its efficiency[11–14].

Reviewers are typically protected by anonymity, and are not rewarded for an accurate and fair job nor held accountable for a sloppy or biased one. Reviewers are thus under little incentive to act in the best interest of science as opposed to their own best interest. Thurner and Hanel[11] showed that if even a small percentage of referees adopt "rational" (i.e. selfish) behavior, rejecting manuscripts they perceive to threaten the visibility of their own work, the average quality of published research suffers substantially. Relative to a scenario where all referees are fully impartial, a percentage as low as 10% of selfish referees in the pool will lower the quality of published work by almost one full standard deviation of the quality of submitted manuscripts.

Although striking, Thurner and Hanel's results[11] are due in part to the extreme behavior adopted by both impartial and selfish referees in their model: impartial referees are constantly raising the bar on manuscripts they will accept, while selfish referees accept manuscripts of below-average quality. A scenario with only impartial referees quickly results in no papers getting published anymore because the standards are so high; in contrast, a scenario with only selfish referees results in the average published paper being of lower quality than the average submitted manuscript. Both outcomes seem too radical.

A second limitation of Thurner and Hanel's model[11] is that it attributes no moderating power to editors, who in reality are expected to buffer poor reviewing in several ways. For example, editors may match manuscripts to suitable referees, aiming to avoid bias caused by conflicts of interest and cliques[7]. Editors may also send manuscripts to multiple referees; Bianchi et al.[9] showed that this also reduces bias, albeit at the expense of more resources invested by scientists on reviewing rather than researching. Editors may bypass reviewers altogether when manuscripts are exceptionally good or bad, accepting or rejecting without review as the case may be[15]. Editors also typically give authors the chance to revise their manuscripts to address reviewers' comments, and may even blacklist referees with a suspicious reviewing record.

Finally, peer review must balance its goal of guaranteeing quality of published work with other desired outcomes such as keeping rejection of good papers at a minimum and distributing the reviewing load evenly across scientists[7,16]. A strict focus on quality of accepted papers may miss other aspects of performance by the peer review process.

Here we extend Thurner and Hanel's agent-based model[11] of peer review to examine how their results change when referees adopt less extreme behavior, such as when impartial referees adopt fixed quality standards and selfish referees care enough about quality control as to reject below-average papers. We then ask to what degree the abovementioned editorial strategies can mitigate the negative impact of selfish referees on the average quality of published work. We note that other authors have examined the effects of some of these editorial strategies[9,13], but this study is the first to place them in the context of different types of referee behavior and compare their effects. We also investigate how editors and referees affect the proportion of

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3005v3 | | rec: 14 Jun 2017, publ: 14 Jun 2017

above-average manuscripts that go unpublished, and the distribution of reviewing effort among all scientists. By examining what aspect of bad reviews most undercuts quality and what editors can do about it, our paper is a step in addressing the need to counter peer review's alarming sensitivity to referee misbehavior.

Methods

The peer review process

Our model is similar to Thurner and Hanel's[11], except that we introduce the editor as an agent in the peer review process. We assume a fixed pool of $N = 1,000$ scientists who variously act as authors and referees. Every year there are two peer review cycles, in each of which half of the scientists submit manuscripts for review. Scientists remain active for 35 years (70 review cycles) before retiring, and are then replaced by new scientists. Each manuscript is handled by one editor and two randomly selected referees. Each manuscript is authored by a single scientist, and authors do not review their own work. We run the simulation for 500 review cycles.

Manuscripts differ by "quality", here abstractly defined as a real number which referees and editors can accurately assess, and which the peer review system aims to maximize in accepted works. Authors in turn differ by "proficiency", here defined as the average quality of their manuscripts. We assume a normal distribution of proficiency with mean μ_a and standard deviation σ_a. Manuscripts authored by an author with proficiency Q fall on a normal distribution of quality with mean Q and standard deviation σ_m. Following Thurner and Hanel, we set $\mu_a = 100$, $\sigma_a = 10$, $\sigma_m = 5$. We coded our model and performed all simulations in the R language[17].

Referee behavior

Impartial referees base their review strictly on the quality of the manuscript. Selfish referees also consider how the manuscript compares with their own research. Like impartial referees, selfish referees impose a minimum quality threshold. Unlike impartial referees, they also impose an upper limit, namely the average quality of their own scientific output. A selfish referee with proficiency Q will therefore reject any paper with quality greater than Q.

We consider two types of impartial referees and two types of selfish referees:

Fixed-standard impartial referees accept manuscripts meeting a fixed minimum standard $Q_{\min}$, and reject all other manuscripts. For concreteness we define this threshold as the 90th percentile of quality distribution across submitted manuscripts, which in our model is a fixed quantity, $Q_{\min} = \mu_a + 1.28\sqrt{\sigma_a^2 + \sigma_m^2} = 114.3$.

Moving-standard impartial referees are similar, except their minimum cutoff is based on the average quality of accepted papers in the previous review cycle, and as such is continually updated. Following[11], we define it as $Q_{\min}(t) = \lambda\, Q_{\min}(t-1) + (1-\lambda)\,\overline{Q}_{\text{accepted}}(t-1)$, where t is the index of the current review cycle, and λ is a fixed parameter controlling how quickly the moving standard asymptotes. We set $\lambda = 0.1$.

Conscientious selfish referees share their minimum standards with fixed-standard impartial referees, $Q_{\min} = 114.3$. The only difference is that they are selfish, as defined above.

Indifferent selfish referees are less particular about the quality of the manuscripts they review, and will accept below-average manuscripts down to $\varepsilon\,\overline{Q}_{\text{submitted}}$, where $\varepsilon < 1$. For consistency with Thurner and Hanel we set $\varepsilon = 0.9$, leading to a minimum cutoff $Q_{\min} = 90$.

For a summary of the minimum and maximum cutoffs of our different types of referees, see Table 1.

Table 1. Minimum and maximum cutoffs imposed by each type of referee. Impartial referees have no maximum cutoff, and accept any manuscript above their minimum standard. The two types differ by whether that minimum is fixed or moving as the review process proceeds. A selfish referee whose average work quality is Q will reject any manuscript above that mark. Indifferent selfish referees care little about quality of published research, and accept manuscripts of below-average quality. In contrast, conscientious selfish referees share minimum standards with fixed-standard impartial referees. $\overline{Q}_{\text{accepted}}$ asymptotes to ≈ 141 when the referee pool consists almost entirely of moving-standard impartial referees, and to gradually lower values as more selfish referees are added.

Referee type	Minimum cutoff	Maximum cutoff
Impartial, moving-standard	$\overline{Q}_{\text{accepted}}(t-1)$	∞
Impartial, fixed-standard	114.3	∞
Selfish, conscientious	114.3	Q
Selfish, indifferent	90	Q

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3005v3 | CC BY 4.0 Open Access | rec: 14 Jun 2017, publ: 14 Jun 2017

Editor strategies

In Thurner and Hanel's model[11], the editor accepts a manuscript if both referees accept it, rejects it if both referees reject it, and accepts it with probability 50% if the referees disagree. In our model, the editor strives to upkeep the quality of published papers by adopting one the following strategies:

Bypassing reviewers. Editors automatically reject a manuscript without review if it fails to meet a minimum standard of quality (here we choose $Q = 101$). This is commonly done by high-profile journals such as Science and Nature. Similarly, editors accept without review manuscripts deemed to be of exceptional quality. For this threshold we use the 90th percentile of papers accepted in the previous round.

Consulting tiebreaking referees. If one referee accepts while the other rejects, the editor sends the manuscript for review by a third referee, whose recommendation breaks the tie.

Returning papers to authors for revision. If at least one referee rejects the manuscript, the editor sends it back to the author for revision and gives it another round of reviewing by the same referees. In our model, the quality increment of the revised manuscript follows a half-normal distribution with mean $\sqrt{2/\pi}\,\sigma_{\mathrm{m}} \simeq 4$.

Matching manuscript quality to referees. Aware that some referees tend to selfishly reject better papers than their own, editors preferentially match manuscripts to referees with somewhat better scientific output. Our editors will pair a manuscript of quality Q with two referees drawn from the pool based on their proficiency Q_{ref} as authors, with the difference $Q_{\mathrm{ref}} - Q$ being normally distributed with mean 5 and standard deviation 5.

Blacklisting selfish referees. If a referee's record of reviews indicates that the referee is selfish, the editor removes that person from the list of referees and no longer sends papers for them to review. Because two impartial referees will never disagree in their reviews but an impartial and a selfish referee or two selfish referees might, a high proportion of disagreements in someone's reviewing record suggests selfishness. A referee is blacklisted when the probability of selfishness given their reviewing record exceeds a threshold p_0. See the supplementary information for the mathematical details.

For simplicity, we assume all journals have the same quality standards and all editors adopt the same editorial strategy in each of our scenarios. In the blacklisting scenario, a referee's reviewing records is shared among all editors. This overestimates the impact of blacklisting, as in reality each journal is likely to keep their own separate record, and therefore referees blacklisted from one journal can still be asked to review manuscripts for other journals.

Results

Impact of referee behavior

When all scientists act as moving-standards impartial referees, the average quality of published papers approaches the very top percentile of submitted manuscripts (Figure 1A). Only the very best manuscripts are accepted, with a rejection rate close to 100% (Figure 1B), with around 50% of rejected papers being of higher quality than the average submitted manuscript (Figure 1C). By comparison, fixed-standards impartial referees keep published research at the top 95 percentile of submitted work and reject 90% of manuscripts, about 45% of which are of above-average quality.

The introduction of indifferent selfish referees, which will accept papers of below-average quality (down to $Q = 90$ given our parameter choices), has a strong impact on the quality of published papers and on rejection rates. As the percentage of indifferent selfish referees increases from near 0 to 90%, mean quality of published work falls from near the 100th percentile of submitted manuscripts to the 50th percentile—thus indistinguishable from a fully random review process. The drop is similar whether the impartial referees have moving or fixed standards (compare black and red curves on Figure 1A).

In stark contrast, the quality of published papers does not suffer nearly as much when selfish referees are conscientious, that is when they share the impartial referees minimum quality standards. As more referees in the pool are selfish but conscientious, mean published quality drops very little or none at all, depending on whether impartial referees have moving or fixed standards (green and blue curves on Figure 1A).

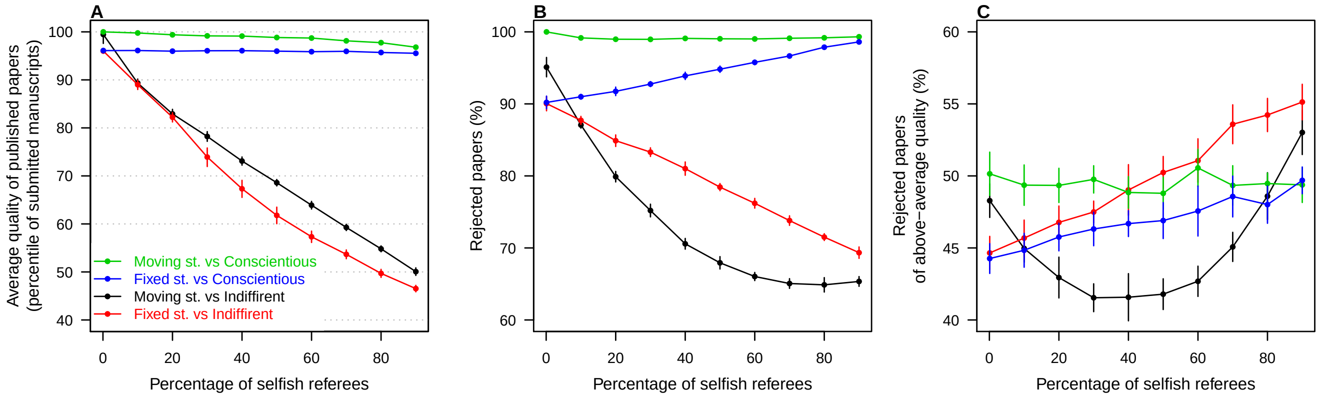
Overall rejection rates are relatively low when indifferent selfish referees are common, and high otherwise (Figure 1B). Given our parameter choices, impartial referees of both kinds and conscientious selfish referees will all tend to reject at least 90% of manuscripts. Indifferent selfish referees have the greatest impact on rejection rates because of their tolerance for low-quality papers.

As selfish referees become more common, so does rejection of good papers (Figure 1C). Comparing Figures 1B and 1C, we notice that as indifferent selfish referees become more common, overall manuscript rejection declines but the percentage of rejected papers with above-average quality increases. When selfish referees are conscientious, above-average papers always constitute less than 50% of rejections. But when they are indifferent and dominate the referee pool, more than one in two rejected papers may be in the upper half of submissions.

These results show that the big impact caused by selfish referees as seen by Thurner and Hanel[11] is due not to their selfishness but to their indifference. That is, they lower the quality of accepted literature not because they reject papers whose

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3005v3 | CC BY 4.0 Open Access | rec: 14 Jun 2017, publ: 14 Jun 2017

Figure 1. Effect of referee behavior. **A**: Average quality of published papers, in percentile of the quality distribution across submitted manuscripts, when impartial and selfish referees are modeled in different ways. Impartial referees have either *moving standards* or *fixed standards*, and selfish referees are either *conscientious* or *indifferent* (see main text for detailed description). Curves show the decline of published quality as the percentage of selfish referees in the pool increases, for each of the four combinations with one type of impartial referee and one type of selfish referee. Thurner and Hanel's results[11] correspond to the black curve. **B**: Percentage of rejected manuscripts by percentage of selfish referees in the pool, for each of the four scenarios. Error bars represent one standard deviation across ten replicates.



quality is too high, but because they accept papers whose quality is too low. Notice that whether impartial reviewers have fixed or moving standards is relatively inconsequential.

Because of the high impact caused by indifference compared to other referee behaviors, we are going to focus on it for the remainder of our study. In the following section, selfish referees are assumed indifferent and impartial referees are assumed to have moving standards, similar to Thurner and Hanel[11].

Impact of editors

None of the editorial strategies listed above could fully neutralize the decline in quality of published research caused by indifferent selfish referees (Figure 2A). However, some of these strategies were able to mitigate the impact.

Submitting contentious manuscripts to tiebreaking reviewers can raise the quality of accepted papers by as much as 10 percentile points of the distribution of submitted manuscripts (Figure 2A). Relative to no editorial action, it can represent an improvement of up to 50%, and the effect is positive as long as selfish referees do not comprise much more than half of the referee pool (Figure 2B). This makes sense, as a tiebreaking reviewer is more likely to be impartial than selfish in those circumstances. As selfish referees become so numerous as to predominate, the tiebreaking strategy becomes counterproductive. Tiebreaking also leads to more manuscript rejections (Figure 2C), but fewer of the rejected manuscripts are of above-average quality (Figure 2D). Again, this holds provided selfish referees are not the majority.

Blacklisting reviewers with a suspiciously high record of disagreements with other reviewers also improves quality of publications. The lower the probability threshold for blacklisting, the bigger the improvement (Figures 2A and 2B). Its effects on rejection statistics are similar to those of tiebreaking (Figures 2B and 2C). However, those benefits are paid for by higher inequality in the referee load across scientists: as some are excluded from the referee list, others must pick up the slack and serve as referees more frequently than they otherwise would. This is reflected as an increase in the Gini coefficient of the distribution of referee load (Figure 3). The impact is higher for a higher proportion of selfish referees in the pool, and for a lower blacklisting threshold.

Other editor strategies did not raise the quality of published papers relative to no action. Those included sidestepping/overruling referees for manuscripts of exceptionally high or low quality, allowing authors to revise their rejected manuscripts for a second round of revision by the same referees, and pairing submitted manuscripts with referees who are unlikely to selfishly reject them.

Bypassing reviewers in the case of exceptional papers has almost no effect except when most referees are selfish (Figures 2A and 2C). This is expected, as most papers are not of exceptional quality. It is perhaps more surprising that even automatic rejection of below-average manuscripts have little effect. However, considering that the standards of our impartial referees are so far above mere mediocrity (they only accept manuscripts above the 90th percentile), it could be that most of the impact of indifferent selfish referees derives from their acceptance of papers between the 50th and 90th percentile, rather than from their acceptance of work below the 50th percentile. In that case, editors keeping out only the lower 50th percentile are leaving most

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3005v3 | CC BY 4.0 Open Access | rec: 14 Jun 2017, publ: 14 Jun 2017

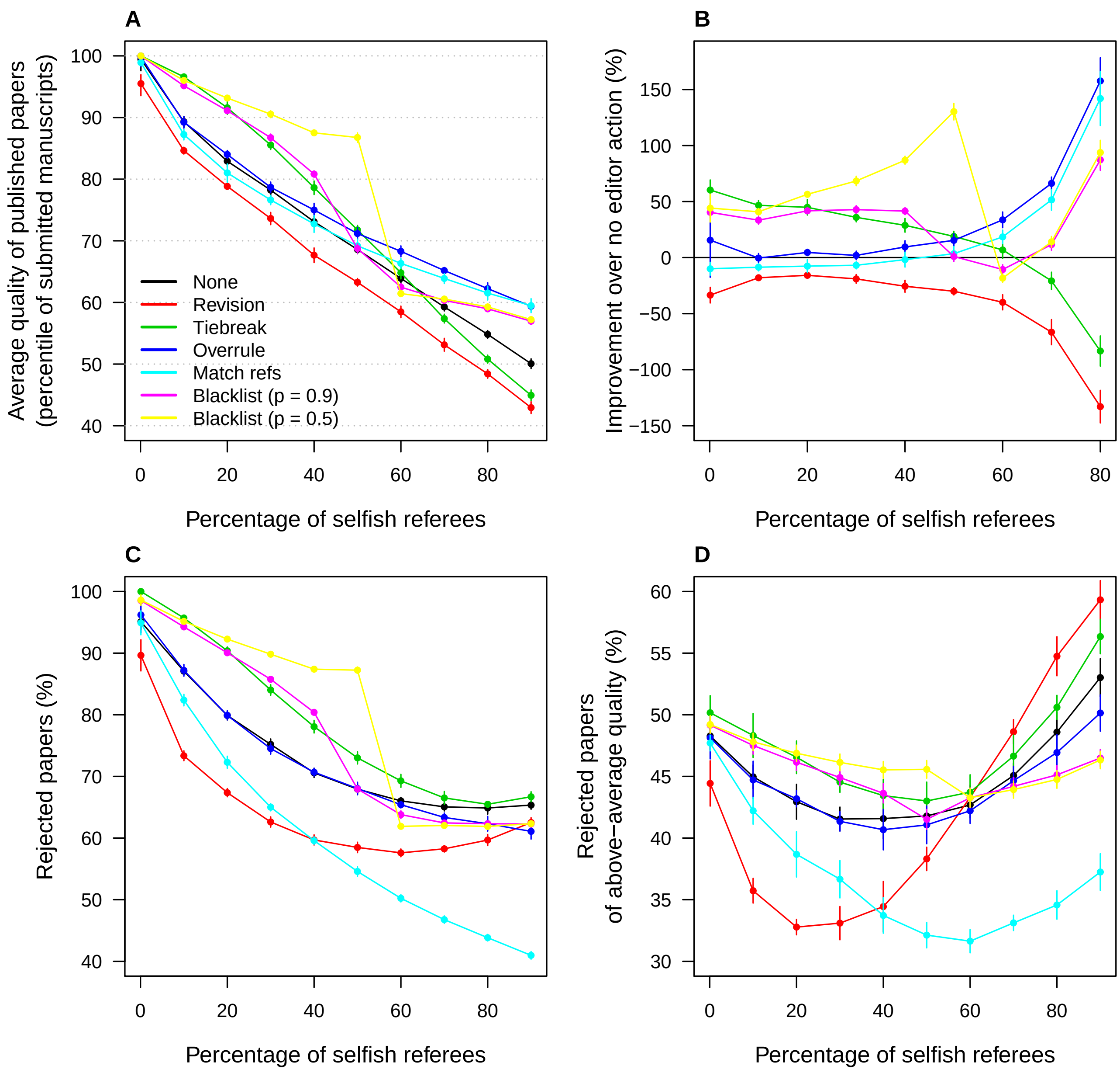


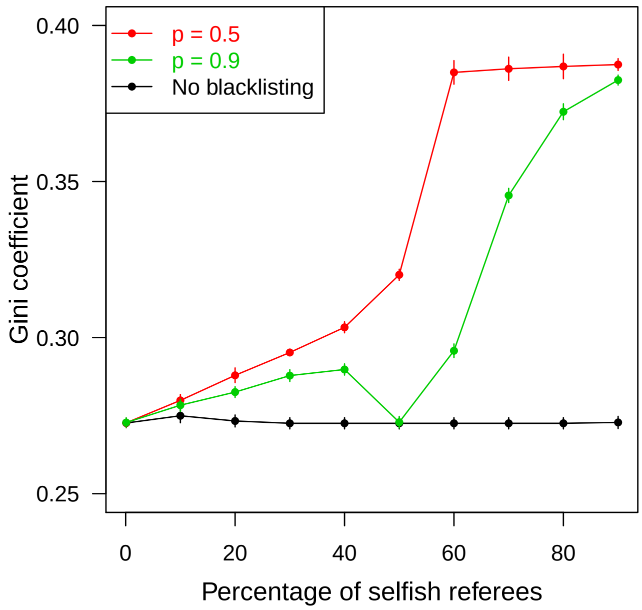
Figure 2. Effect of editor strategies. A: Average quality of accepted papers by percentage of selfish referees in the pool, measured in terms of the percentile of the quality distribution across submitted manuscripts. *None*: no editorial action; *Revision*: submitted papers with at least one rejection are sent back for revision and given a second round of reviews; *Tiebreak*: if the paper is accepted by one referee but rejected by the other, the editor sends it to a third referee; *Overrule*: papers of exceptionally poor/good quality are automatically rejected/accepted by the editor without review; *Matching referees*: editor pairs submitted manuscripts with editors of higher quality to avoid biased rejection from selfish referees; *Blacklist*: editor blacklists referees with a higher probability of being selfish than a threshold p. Shown are results for $p = 0.5,\ 0.9$. **B**: Percentage improvement on average quality of published papers relative to no action. **C**: Percentage of rejected manuscripts by percentage of selfish referees in the pool. Color scheme in (**B**) and (**C**) identical to (**A**).

of the problem unaddressed.

Matching manuscripts to editors with a superior research output minimizes rejection based on a sense of threat or competition. This leads to much lower rejection rates, both in general (Figure 2C) and of above-average papers (Figure 2D). However, it does little to improve the quality of those papers that do get accepted, unless selfish referees vastly outnumber impartial ones

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3005v3 | CC BY 4.0 Open Access | rec: 14 Jun 2017, publ: 14 Jun 2017

Figure 3. Evenness in referee load. Gini coefficient of the distribution of referee load plotted against percentage of selfish referees in the pool. Red and green curves show scenarios when editors blacklist referees with a higher than 50% and 90% chance of being selfish, respectively. Black curve shows Gini coefficient when editors do not blacklist. Referee load is defined as the number of times a scientist served as a referee. The Gini coefficient is half the average absolute difference in referee load across scientists, scaled by the average referee load of all scientists.



(Figure 2A and 2B). This is because this strategy does not address the aspect of referee behavior that most impacts quality of published work, namely accepting low-quality papers.

Sending rejected manuscripts back for revision by the authors, a very common practice in many journals, does not help raise published quality either. In fact it lowers it, whether selfish referees are common or rare (Figures 2A and 2B). We explain this counterintuitive result thus: two indifferent selfish referees will only disagree if the manuscript is too good for one of them. Improving the quality of the manuscript will therefore not change the verdict of that referee, and might even make it too good for the other referee as well. Revising in this case does not help. An impartial and an indifferent selfish referee will disagree if the manuscript is either not good enough for the impartial referee or too good for the selfish referee. In the latter case revising does not change either review, and in the former it may work for the author, or it may overshoot the acceptance window of the selfish referee. Of course the original manuscript can be so bad that both referees reject it. In that case, revising may turn one or both reviewers around and get the paper in. Overall, revising is likely to make the most difference for papers of marginally acceptable quality, and their publication does little to boost the average quality of published work. More importantly, it does nothing to stop the acceptance of low quality work. What it does accomplish is a substantial decline in rejection of above-average manuscripts, particularly when selfish referees do not dominate the pool entirely (Figure 2D).

Discussion

Cheating referees severely undercut the quality of published literature[11]. We have shown that the brunt of this impact is due to their willingness to accept low-quality manuscripts, as opposed to their tendency to reject high-quality manuscripts perceived as competition. Published quality will remain high as long as referees commit to rejecting papers that fail to meet minimum standards set by journals, regardless of their potential inclination to sabotage work better than their own. Self-interested sabotaging has a relatively stronger impact on rejection rates, both in general and of above-average manuscripts.

We found that editors can mitigate, but not neutralize, the impact of indifference to quality by forwarding disputed manuscripts to tiebreaking referees. Our results corroborate previous findings that sending manuscripts to multiple reviewers helps to protect the quality of the process. However, a blanket strategy of using multiple reviewers has systemic costs in terms of time diverted from research activities towards reviewing[9]. By comparison, the tiebreaking strategy keeps such costs to a minimum. On the other hand, we note that effective strategies for maximizing quality also tend to increase the proportion of

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3005v3 | CC BY 4.0 Open Access | rec: 14 Jun 2017, publ: 14 Jun 2017

rejected manuscripts with above-average quality, especially if reviewers continually raise their standards.

Removing reviewers with a suspicious record from the referee pool increases quality. Grimaldo and Paolucci[13] found that blacklisting improves the quality of published research, minimizes disagreement between referees, and reduces rejection of good papers. However, they assumed that selfish referees share their concern for quality with impartial referees, and they only considered rejected papers that would have been accepted in the absence of selfish referees. Our results are consistent with Grimaldo and Paolucci's[13], but we also found that blacklisting is less successful if selfish reviewers accept low-quality manuscripts, and we showed that under a broader definition of good papers that includes all manuscripts of above-average quality, blacklisting actually increases rejection of good papers. Furthermore, we emphasized that this editorial strategy increases the load on whitelisted referees to compensate for blacklisted referees. Because overburdened reviewers have less time for their own research[7], this ultimately punishes good referees.

Pairing manuscripts with good-quality reviewers and allowing authors to revise rejected manuscripts failed to improve quality relative to no action, although both strategies substantially lowered the proportion of above-average rejected manuscripts. Automatic acceptance or rejection without review of exceptional manuscripts had little impact on both mean published quality and rejection rates, although it guarantees that the very best papers will be published and the very worst will not.

Our study made several simplifying assumptions. We assumed no referee bias other than a selfish interest in sabotaging competition. In reality, reviewers' judgment may be colored by their opinion of the authors. (This has led some journals to adopt a double-blind reviewing process, e.g. Nature, The American Naturalist, Social Science & Medicine). We also assumed editors have no bias of their own, although editor bias can have an even bigger impact on peer review than referee bias[10]. Differences in quality standards across different journals also affect the outcome of peer review models[13], but were left out of our study so we could focus on the impact of editorial strategies. Finally, we assumed that manuscripts are judged based on a single agreed-upon quantity that captures their scientific merit and publication worth. In reality, manuscripts are evaluated on several independent measures (novelty, clarity, accuracy, relevance), and earnest impartial reviewers and editors may disagree over the merits of any given manuscript. The latter will contribute further noise to the process, but no systematic bias. More sophisticated models accounting for these omissions should reach the same qualitative conclusions regarding the impact of referee indifference and the mitigating potential of editor strategies.

We conclude that peer review in its current format offers little incentive for altruistic behavior from referees, and has limited tools to safeguard the efficiency of the process. Efforts to minimize the number of bad papers accepted must balance the simultaneous goals of also minimizing the number of good papers rejected, and evenly distributing the burden of referee service among all scientists. Editorial mitigating strategies help, but a structural change that rewards good reviewing practices and discourages cheating may have a stronger impact. Finally, we note that several modifications to the classical peer review process have been proposed. Recent tests indicate that alternatives such as bidding and review-sharing may outperform classical peer review in speed of publication and quality control[15, 18]. It remains to be seen whether such alternatives are more robust to bias and cheating referee behavior.

Acknowledgements

J.OD. acknowledges the Simons Foundation Grant #376199 and McDonnell Foundation Grant #220020439.

Author contributions statement

R.D. and J.OD. conceived the study. R.D. wrote the code and analysed the results. All authors reviewed the manuscript.

Additional information

Competing financial interests

The authors declare no competing financial interests.

References

1. Ware, M. Peer review: benefits, perceptions and alternatives. *Publ. Res. Consortium* 20 (2008).
2. Jefferson, T. O., Alderson, P., Davidoff, F. & Wager, E. Editorial peer-review for improving the quality of reports of biomedical studies (Cochrane Methodology Review). *Cochrane Database Syst. Rev.* (2007). URL http://www.update-software.com/cochrane/MR000016.pdf. DOI 10.1002/14651858.MR000016.pub3.www.cochranelibrary.com.
3. Mahoney, M. J. Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cogn. Ther. Res.* **1**, 161–175 (1977). DOI 10.1007/BF01173636.

4. Rothwell, P. M. & Martyn, C. N. Reproducibility of peer review in clinical neuroscience. Is agreement between reviewers any greater than would be expected by chance alone? *Brain* **123 (Pt 9**, 1964–1969 (2000). DOI 10.1093/brain/123.9.1964.

5. Jefferson, T., Alderson, P., Wager, E. & Davidoff, F. Effects of editorial peer review: A systematic review. *J. Am. Med. Assoc.* **287**, 2784–2786 (2002). URL http://dx.doi.org/10.1001/jama.287.21.2784{%}5Cnhttp://jama.jamanetwork.com/data/Journals/JAMA/4834/JRV10095.pdf. DOI 10.1001/jama.287.21.2784.

6. Roebber, P. J. & Schultz, D. M. Peer review, program officers and science funding. *PLoS ONE* **6**, 2–7 (2011). DOI 10.1371/journal.pone.0018680.

7. Cabotà, J., Grimaldo, F. & Squazzoni, F. Do editors have a silver bullet? An agent-based model of peer review. *Proc. 28th Eur. Conf. on Model. Simul. ECMS 2014* 725–731 (2014). URL http://www.scs-europe.net/dlib/2014/ecms14papers/fes{_}ECMS2014{_}0063.pdf.

8. Thorngate, W. & Chowdhury, W. By the Numbers: Track Record, Flawed Reviews, Journal Space, and the Fate of Talented Authors. In Kaminsky, B. & Koloch, G. (eds.) *Advances in Social Simulation. Advances in Intelligent Systems and Computing*, vol. 229, 177–188 (Springer, Berlin, Heidelberg, 2014). URL http://link.springer.com/10.1007/978-3-642-39829-2. DOI 10.1007/978-3-642-39829-2. 9780201398298.

9. Bianchi, F. & Squazzoni, F. Is three better than one? Simulating the effect of reviewer selection and behavior on the quality and efficiency of peer review. *Proc. 2015 Winter Simul. Conf. (WSC). IEEE, 2015* 779–790 (2015).

10. Wang, W. *et al.* Editorial behaviors in peer review. *SpringerPlus* **5**, 903 (2016). URL http://springerplus.springeropen.com/articles/10.1186/s40064-016-2601-y. DOI 10.1186/s40064-016-2601-y.

11. Thurner, S. & Hanel, R. Peer-review in a world with rational scientists: Toward selection of the average. *Eur. Phys. J. B* **84**, 707–711 (2011). DOI 10.1140/epjb/e2011-20545-7. 1008.4324.

12. Squazzoni, F. & Gandelli, C. Saint Matthew strikes again: An agent-based model of peer review and the scientific community structure. *J. Informetrics* **6**, 265–275 (2012). URL http://dx.doi.org/10.1016/j.joi.2011.12.005. DOI 10.1016/j.joi.2011.12.005.

13. Grimaldo, F. & Paolucci, M. A Simulation of Disagreement for Control of Rational Cheating in Peer Review. *Adv. Complex Syst.* **16**, 1350004 (2013). URL http://www.scopus.com/inward/record.url?eid=2-s2.0-84879650706{&}partnerID=tZOtx3y1. DOI 10.1142/S0219525913500045.

14. Paolucci, M. & Grimaldo, F. Mechanism change in a simulation of peer review: From junk support to elitism. *Sci.* **99**, 663–688 (2014). DOI 10.1007/s11192-014-1239-1.

15. Allesina, S. Modeling peer review: an agent-based approach. *Ideas Ecol. Evol.* **5**, 27–35 (2012). URL http://library.queensu.ca/ojs/index.php/IEE/article/view/4447. DOI 10.4033/iee.2012.5b.8.f.

16. Squazzoni, F. & Gandelli, C. Opening the Black-Box of Peer Review : An Agent-Based Model of Scientist Behaviour. *J. Artif. Soc. Soc. Simul.* **16**, 1–8 (2013).

17. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2017). URL https://www.r-project.org/.

18. Kovanis, M., Trinquart, L., Ravaud, P. & Porcher, R. Evaluating alternative systems of peer review: a large-scale agent-based modelling approach to scientific publication. *Sci.* 1–21 (2017). DOI 10.1007/s11192-017-2375-1.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3005v3 | CC BY 4.0 Open Access | rec: 14 Jun 2017, publ: 14 Jun 2017