Title: Cooperation between referees and authors increases peer review accuracy

Authors: Jeffrey T. Leek*[1], Margaret A. Taub[1], Fernando J. Pineda[1,2]

1. Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205
2. Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21025

Abstract:

Peer review is fundamentally a cooperative process between scientists in a community who agree to review each other's work in an unbiased fashion. Peer review is the foundation for decisions concerning publication in journals, awarding of grants, and academic promotion. Here we perform a laboratory study of open and closed peer review based on an online game. We show that when reviewer behavior was made public under open review, reviewers were rewarded for refereeing and formed significantly more cooperative interactions (13% increase in cooperation, P = 0.018). We also show that referees and authors who participated in cooperative interactions had an 11% higher reviewing accuracy rate (P = 0.016). Our results suggest that increasing cooperation in the peer review process can lead to a decreased risk of reviewing errors.

Text:

Peer review is the foundation for decisions concerning publication in journals, awarding of grants, and academic promotion. However, the anonymous nature of peer review is increasingly under scrutiny [1,2,3,4], and some journals have

considered or already moved to open peer review [2,5,6,7]. Despite its central role in the scientific process, the underlying social dynamics and accuracy of peer review under alternative systems are difficult to study. We have performed experiments to collect quantitative data about the social behavior of referees in anonymous (closed) and non-anonymous (open) peer review using an online game.

Our model system for peer review was an online game launched from the Amazon EC2 cloud played by 7 – 10 individuals over a fixed period. Players were graduate students, postdoctoral fellows, research scientists or principal investigators, all of whom are members of a single research laboratory. The game was designed to replicate several components of editorial peer review: 1. most reviewers know the authors of the papers they referee, 2. peer review is usually performed within relatively small communities of individuals[8], and 3. peer review involves repeated interactions between referees and authors. The game's interface presented players with multiple-choice questions similar to those found on the Graduate Record Exam (GRE)[9]. At any point in the game a player chose between solving problems or reviewing (accepting or rejecting) solutions submitted by other players. The software also played the role of journal editor and randomly assigned submitted solutions to players for review. At the end of the game, the two players with the largest number of accepted submissions received monetary rewards, reflecting the conventional "publish or perish" academic incentives.

Individual games were played in either a closed mode, or in an open mode. In the closed mode, the reviewers were anonymous (**Supplementary Fig 1 left column**). In the open mode, players knew which reviewer accepted or rejected each of their submissions (**Supplementary Fig 1 right column**). The public information under the closed mode was the number of submitted solutions that were accepted. In the open mode, both the number of submitted solutions that were accepted and the number of times each player accepted a peer's solution were public.

We recruited members of six research laboratories at Johns Hopkins University to play the Peer Review Game in closed mode (3 labs, n=8, 8, and 9 players) and open mode (3 labs, n=7,10, and 8 players). Each laboratory played the game for T = 40 minutes. We collected a total of 1,143 solutions and 666 reviews over the course of the six experiments. In the open system each solution a player accepted led to an increased probability their own next submission would be accepted (2% increase per accepted solution, P = 0.047) – which agreed with our theoretical analysis (**Supplementary Material**). In previous studies of iterated games, costly punishment has been shown to be negatively associated with payoff – in other words "winners don't punish"[10]. Under the open system one of the top two reviewers was always one of the winners of the game, suggesting that reviewers were rewarded for their good behavior toward other players (**Supplementary Materials**).

An important question is whether making reviewing behavior public facilitates cooperation. For each experiment we calculated a pair-wise measure of cooperation between players (**Supplementary Materials**). The open review experiments showed more cooperative connections than the closed experiments (22% versus 9% respectively, P = 0.018, **Fig 1**). It was not immediately clear that cooperation between referees and authors would increase reviewing accuracy. Intuitively, one might expect that players who cooperate would always accept each other's solutions - regardless of whether they were correct. However, we observed that when a submitter and reviewer acted cooperatively, reviewing accuracy actually increased by 11% (P = 0.016). The difference in accuracy was significant even after adjusting for the fact that some solvers had higher accuracy than others (11% increase in accuracy, P = 0.039). The increase in reviewing accuracy was mediated by cooperative interactions between players, since overall accuracy was comparable under open and closed peer review (1% more accuracy under closed, P = 0.762).

We have shown that one mechanism for increased cooperation is making reviewer information public, but other mechanisms for improving cooperation in the review process, such as reducing calls for unnecessary experiments[11] or through appropriate incentives to referees[12] could also be explored. Regardless of mechanism, in this era of increasing competition for publication and grants, cooperation is vital for accurate evaluation of scientific research.

Acknowledgements

We would like to acknowledge a Johns Hopkins Bloomberg School of Public Health Faculty Innovation Award to J. Leek for funding.

Informed Consent

Written informed consent was obtained from all participants in this study. This specific study was approved by the Johns Hopkins Bloomberg School of Public Health IRB with project number 3316.

References

1. Fabiato A (1994) Anonymity of reviewers. Cardiovasc Res 28: 1134-1139; discussion 1140-1135.
2. Smith R (1999) Opening up BMJ peer review - A beginning that should lead to complete transparency. British Medical Journal 318: 4-5.
3. Wenneras C, Wold A (1997) Nepotism and sexism in peer-review. Nature 387: 341-343.
4. Mietchen D (2011) Peer reviews: make them public. Nature 473: 452-452.
5. Editorial N (2006) Peer Review Debate. Nature Web Focus.
6. Pulverer B (2010) Transparency showcases strength of peer review. Nature 468: 29-31.
7. Pulverer B (2011) Peer reviews: some are already public. Nature 474.
8. Knoll E (1990) The communities of scientists and journal peer review. JAMA 263: 1330-1332.
9. Mathur H.
10. Dreber A, Rand DG, Fudenberg D, Nowak MA (2008) Winners don't punish. Nature 452: 348-351.
11. Ploegh H End the wasteful tyranny of reviewer experiments. Nature 472: 391.
12. Schroter S, Tite L (2007) Why do peer reviewers decline to review? A survey. Journal of Epidemiology and Community Health 61: 9-12.

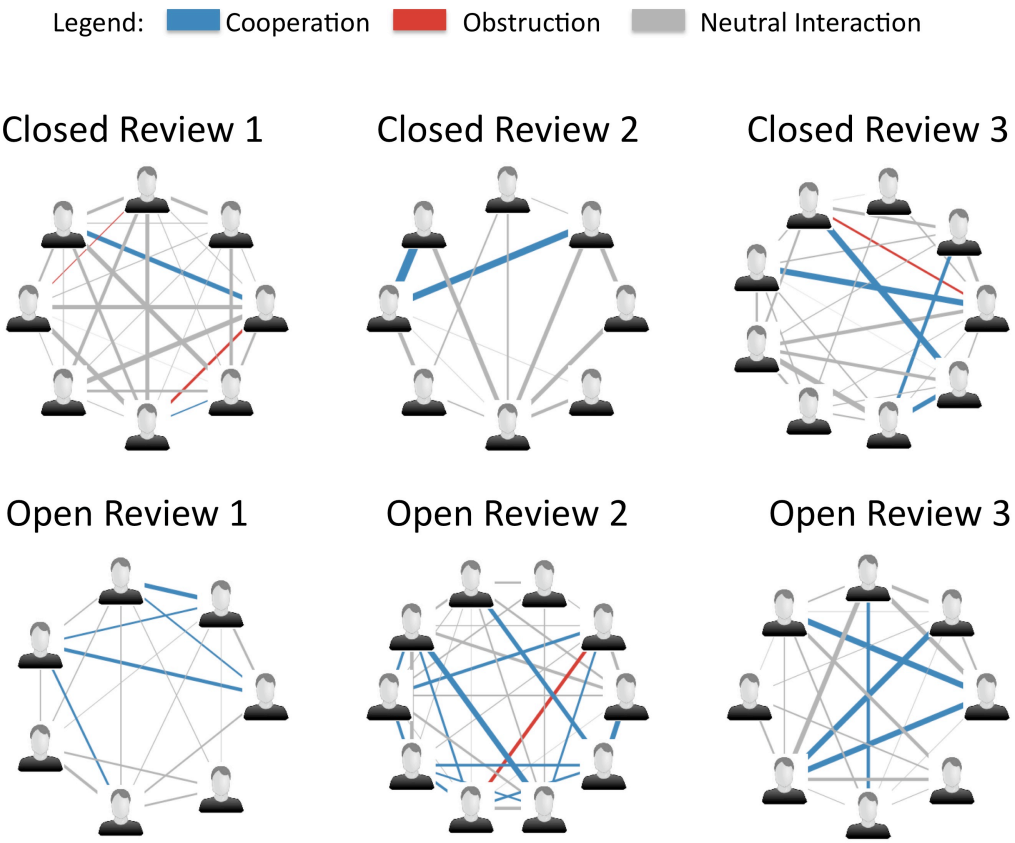

Fig 1 Open reviewers are more cooperative than closed reviewers Each panel shows the cooperation network for one of the peer review experiments. The thickness of the line indicates the amount of interaction and color indicates the type of interaction. Cooperation (blue) is defined as above average probability of both players accepting each other's solutions. Obstruction (red) is defined as below average probability of both players accepting each other's solutions. Under closed review (top row) there is less cooperation between players than under open review (bottom row).