![JASSS logo]

Flaminio Squazzoni and Károly Takács (2011)

Social Simulation That 'Peers into Peer Review'

Abstract

This article suggests to view peer review as a social interaction problem and shows reasons for social simulators to investigate it. Although essential for science, peer review is largely understudied and current attempts to reform it are not supported by scientific evidence. We suggest that there is room for social simulation to fill this gap by spotlighting social mechanisms behind peer review at the microscope and understanding their implications for the science system. In particular, social simulation could help to understand why voluntary peer review works at all, explore the relevance of social sanctions and reputational motives to increase the commitment of agents involved, cast light on the economic cost of this institution for the science system and understand the influence of signals and social networks in determining biases in the reviewing process. Finally, social simulation could help to test policy scenarios to maximise the efficacy and efficiency of various peer review schemes under specific circumstances and for everyone involved.

Keywords:
    Peer Review, Social Simulation, Social Norms, Selection Biases, Science Policy

Why peer review is important for science

1.1   Peer review is one of the most important facets that makes science a complex social system. It became the cornerstone of science from when in 1752 the Royal Society of London obtained the fiscal responsibility for *Philosophical Transactions* and peers were systematically and voluntarily involved to contribute to the quality and excellence of their publications. Now, it is applied to many spheres of scientific activity such as funding, publication, recruitment and even research productivity evaluation. It is essential for institutional agencies in evaluating research grants, for journal and book editors to evaluate the quality of submissions, for scientists to increase the quality of their work, as well as for policy makers to guarantee that taxpayer money is invested in a credible and well functioning system (Squazzoni 2010).

1.2   More importantly, peer review encapsulates the very idea of science that new lines of research are experimentally pursued by scientists through a continuous, decentralised and socially shared trial and error process. It therefore helps science to be self-regulated by determining scientific pay-offs. It directly or indirectly determines how funds and carriers are allocated in science and therefore makes a big difference every day.

1.3   Although peer review can take different forms, it can generally be defined as *a distributed and decentralised mechanism that makes evaluation and improvement of complex scientific products possible through voluntary and impersonal cooperation among peers*. Scientists interact in different roles as journal editors, authors and reviewers. This intensive interaction is guided by a complex set of socially shared norms and values that are the essence of the 'scientific method'. The normative foundations of science include the importance of communalism, universalism, disinterestedness and organised scepticism (Merton 1973), but we can also add the objective search for truth, respect for evidence, tolerance, trust and reputation among peers (Durant and Ibrahim 2011).

What are the problems?

2.1   In our view, there are at least three reasons that require a large scale involvement of social simulation in science and peer review investigation. First, there is evidence that peer review and evaluation in science generally are now under increasing strain. The

tremendous expansion of specific topics, interdisciplinary research and the increasing sophistication of research technologies on one hand and the growing number of journals, conferences and funding agencies on the other, make peer review extremely difficult and largely overexploited (Alberts, Hanson and Kelner 2008). The continuous stratification of the scientific community into a mosaic of specialties and the consequent growth of inter-disciplinary collaboration have increased knowledge asymmetries between editors, reviewers and authors, and so the likelihood of cheating and moral hazards. This has complicated the management of peer review and undermined the possibility of evaluating research proposals and journal submissions appropriately through individually isolated peer review (Grainger 2007).

2.2   A recent survey found that there were approximately 1,346,000 peer-reviewed scientific journal articles published world-wide in 2006, with approximately 70% covered by ISI (Björk, Roos and Lauri 2009). About the same figure was found by Elsevier in an answer to a UK House of Commons' committee in 2004. As the scientific journal publishing market is estimated to growing steadily at about 3.5% annually since the 1970s (estimation on 2001, Mabe and Amin 2001), we can realise the over-exploitation of this important mechanism, not to mention the case of books, research grants, universities and research institutes' productivity evaluation where peer review is also involved. Moreover, the expected world-wide convergence towards an Anglo-American competitive model to allocate resources in science (particularly for funding) and the increase in detail and spheres where peer evaluation is presumably massively applied, will increase even further its present exploitation. In short, there are *strong reasons to doubt that voluntary, uncompensated peer review can go on efficiently* bearing it present burden without reform. Therefore, investigating peer review is fundamental to understand how to exploit this important mechanism more efficiently without deteriorating it.

2.3   Secondly, if some reform is needed, this should follow scientific evidence. It is therefore frustrating to see that peer review mechanisms are dramatically under-investigated (e.g., Kassirer and Campion 1994; Horrobin 2001; Smith 2006). There are anecdotes, personal memories of journal editors and rare investigations on specific cases (e.g., Alberts, Hanson and Kelner 2008; Lamont 2009; Pulvener 2010), but no robust experimental and theoretical knowledge. It seems that scientists devote extended efforts to investigate everything except those particular evaluation mechanisms that make science what it is.

2.4   Of course, one may say that this is not a problem. As a matter of fact, over centuries of evolution of science, we have cumulated experience and discovered practices, standards and technologies that have helped us to guide evaluation and peer review toward efficient and socially shared criteria. Therefore, why bother studying it? The problem is that we have evidence of certain peer review deficiencies in guaranteeing the quality and efficiency of evaluation processes and preventing scientific misconduct (e.g., Bornmann and Daniel 2005; Couzin 2006; Mayo et al. 2006; Nature 2006). Smith (1997; 2006) indicated that peer review is a "black box" with very little knowledge on the benefits or serious evidence on deficiencies. Horrobin (2001) argued that "a process that is central to the scientific endeavour as peer review has no validated experimental base". Therefore, we think that there is room for social simulation to fill this gap by examining peer review, spotlighting social mechanisms behind it at the microscope and understanding their implications for the science system.

2.5   Thirdly, we believe that there is interest in this discussion as many journal editors funding agencies have warned us about the need for revising peer review. Certain attempts to introduce measures to improve the situation have followed trial and error approaches (Squazzoni 2010). Recently, from the authoritative columns of *Science*, Alberts, Hanson and Kelner (2008) have suggested the need for seriously looking at peer review to improve its efficiency and guarantee its sustainability. In short, peer review is also a policy problem which has not yet been seriously addressed scientifically.

🌐   How can social simulation help?

3.1   If this is the case, what can social simulation do, given its focus on modelling social interaction and human behaviour? We think it can do a lot. Among the aspects that are worth investigating, the most important are the following (please, bear in mind that we are sociologists and consequently this list may be biased).

3.2   First, social simulation could help to improve our understanding of why voluntary peer review works at all. Inspired by the literature on the emergence of voluntary cooperation (e.g., Axelrod 1984; 1997) and the emergence of social norms (Axelrod 1986; Coleman 1986), social simulators could study how the seemingly irrational voluntary commitment of editors and reviewers could emerge and investigate the specific norms internalized by peers. The fragility of peer review against amoral behaviour of the agents involved (particularly reviewers) has been recently investigated by a noticeable agent-based model that indicated that even a small fraction of unfair agents can drastically lower the quality of publications (Thurner and Hanel 2010). We think that the well-recognised social simulation literature on cooperation and social norms might help to find measures to increase the strength of norms for more robust cooperation in peer review. We think that the Merton-inspired "middle-range models" capable of combining theoretical intuitions and evidence on well-specified empirical puzzles could be extremely beneficial to understand the peculiarities of peer review mechanisms.

3.3   Secondly, following Hauser and Fehr (2007), social simulation could also help to explore the relevance of social sanctions and reputational motives to increase the commitment of agents involved. On the one hand, the sanction-side of cooperation in peer review—against lazy or unfair reviewers or unreliable authors—is largely unexploited by journals or research funding agencies, while evidence unequivocally demonstrates its crucial relevance (e.g., Gintis 2000). On the other hand, while reputational incentives might guide authors and editors, the current practice of peer review does not allow for reviewer reputation building. As

we know that reputation is one of the most efficient engines of cooperation (e.g., Wedekind and Milinski 2000), it is time to explore what are the consequences of adjustments in reputational benefits provision for reviewers. Obviously, these investigations could also have important policy implications as they could suggest measures for journals to improve reviewer reliability and consequently to enforce author/investigator fairness.

3.4 Thirdly and more generally, social simulation could help to cast light on the economic cost of peer review for the science system. There is interesting literature on the so-called "grant mania" that shows the incredible lost research productivity for the science system. This is due to the time spent in grant applications and reviewing (e.g., Spier 2002; Goldsworty 2009; Gordon and Poulin 2009; Schaffer 2009). Abstract social simulation models could easily look at important aspects of this. Some examples could be the macro consequences of the trade-off between publishing and reviewing, alternative economic resource allocation schemes, such as peer review, bibliometric indexes, equally distributed baseline grants and the specific circumstances one scheme is better than another.

3.5 Another objective for social simulation could be to understand the influence of signals and social networks in determining biases in the reviewing process. We know that the so-called "old-boyism" is strongly affected by the social embeddedness of scientists and we also know that gender and other signals might strongly bias reviewing (e.g., Bornmann and Daniel 2005; Obrecht, Tibelius and D'Aloiso 2007). An interesting example exists where an agent-based model was build to examine the effect of social influence on scientists' behaviour (Martins 2010). However this should be done more precisely for reviewing. Investigations about the social embeddedness of the review process could help to find answers to some basic questions of the philosophy of science. Some examples could be: how can major discoveries break through the "old-boys" barriers? How is the social hierarchy maintained in science and how does it develop over time?

3.6 Finally, social simulation could help to test policy scenarios to maximise the efficacy and efficiency of various peer review schemes under specific circumstances and for everyone involved. For instance, it is reasonable to suppose that journal editors, authors and reviewers have conflicting interests. While editors might be interested in receiving severe judgments from experts to defend the prestige of their journals, submission authors might be interested in receiving a fair treatment, justified judgments and well-detailed reviewer reports that help them to improve the quality of their work (e.g., Schwartz and Zamboanga 2009). These conflicting objectives, far from being simultaneously taken into account, are not often even fairly contemplated presently. If we look at the reviewers' side, there is evidence that reviewing effort follows a Pareto-like distribution, in which a few scientists are responsible for the large majority of submission reviews. A survey conducted in 2007 on a sample of 3,000 scientists showed that most active reviewers covered about 80% of all reviews, with an average of 14 reviews per year (Ware 2007). Therefore, models that can test measures to distribute the reviewing effort more equally without losing reliability and quality, would be welcomed. It is worth noting that these policy analyses could also help to improve the web-platforms currently used by many of us to manage peer review in journals and conferences. This would allow editors to set up reviewing schemes that maximise their objectives and guarantee efficient reviewing.

3.7 This said, given that many datasets already exist (e.g., in journal and conference submissions), a strong recommendation is not only to approach this topic through abstract if important theoretical models, but also to work with empirically grounded models. To do so, it might be essential to collaborate with experts in the field (e.g., economists and science sociologists), and involve relevant stakeholders, such as journal editors, conference chairs and research funding managers, in joint research. Firstly, models are more informative when addressed to well-specified empirical puzzles and grounded on empirical data. Secondly, by involving experts in the field, besides reducing of the risk of reinventing the wheel, social simulation could become accepted in other well-established communities. Thirdly, by collaborating, we could prove that policy analysis can benefit from modelling and understanding social interaction in complex systems, such as science (e.g., Squazzoni and Boero 2010).

References

ALBERTS, B., Hanson, B., and Kelner, K. L. (2008) Reviewing Peer Review. *Science* 321, p. 15. [doi:10.1126/science.1162115]

AXELROD, R. (1984) *The Evolution of Cooperation.* New York, Basic Books.

AXELROD, R. (1986) An Evolutionary Approach to Norms. *American Political Science Review,* 80, pp. 1095-1111. [doi:10.2307/1960858]

AXELROD, R. (1997) *The Complexity of Cooperation. Agent-Based Models of Competition and Collaboration.* Princeton NJ, Princeton University Press. [doi:10.1515/9781400822300]

BJÖRK, B.- C., Roos, A., and Lauri, M. (2009) Scientific Journal Publishing: Yearly Volume and Open Access Availability. *Information Research* 14, 1. http://informationr.net/ir/14-1/paper391.html

BORNMANN, L. and Daniel, H.- D. (2005) Selection of Research Fellowship Recipients by Committee Peer Review. Reliability, Fairness and Predictive Validity of Board of Trustees' Decisions. *Scientometrics* 63, 2, pp. 297-320. [doi:10.1007/s11192-005-0214-2]

COLEMAN, J. S. (1986) Social Structure and the Emergence of Norms among Rational Actors. In: Diekmann, A. and Mitter, P.

(eds.): *Paradoxical Effects of Social Behavior. Essays in Honor of Anatol Rapoport.* Heidelberg, Physica. [doi:10.1007/978-3-642-95874-8_6]

DURANT, J. and Ibrahim, A. (2011) Celebrating the Culture of Science. *Science*, 331, p. 1242. [doi:10.1126/science.1204773]

COUZIN, J. (2006) ... And How the Problems Eluded Peer Reviewers and Editors. *Science* 311, pp. 614-615. [doi:10.1126/science.311.5757.23]

GINTIS, H. (2000) Strong Reciprocity and Human Sociality. *Journal of Theoretical Biology*, 206, pp. 169-179. [doi:10.1006/jtbi.2000.2111]

GOLDSWORTY, J. (2009) Research Grant Mania. *Australian Universities Review* 50, 2, pp.: 17-24.

GORDON, R. and Poulin, B. J. (2009) Cost of the NSERC Science Grant Peer Review System Exceeds the Cost of Giving every Qualified Researcher a Baseline Grant. *Accountability in Research: Policies and Quality Assurance* 16, 1: 13-40. [doi:10.1080/08989620802689821]

GRAINGER, D. W. (2007) Peer Review as Professional Responsibility: A Quality Control System Only as Good as the Participants. *Biomaterials*, 28, pp. 5199-5203. [doi:10.1016/j.biomaterials.2007.07.004]

HAUSER, M. and Fehr, E. (2007) An Incentive Solution to the Peer Review Problem. *PLOS Biology* 5: e107. [doi:10.1371/journal.pbio.0050107]

HORROBIN, D. F. (2001) Something Rotten at the Core of Science? *Trends in Pharmacological Sciences* 22, 2, pp. 51-52. [doi:10.1016/S0165-6147(00)01618-7]

KASSIRER, J. P., Campion E. W. (1994) Peer Review: Crude and Understudied. *Journal of American Medical Association*, 272, pp. 96-97. [doi:10.1001/jama.1994.03520020022005]

LAMONT, M. (2009) *How Professors Think: Inside the Curious World of Academic Judgment.* Cambridge, MA: Harvard University Press. [doi:10.4159/9780674054158]

MABE, M. and Amin, M. (2001) Growth Dynamics of Scholarly and Scientific Journals. *Scientometrics* 51, 1, pp. 147-162. [doi:10.1023/A:1010520913124]

MARTINS, A. (2010) Modeling Scientific Agents for a Better Science. *Advances in Complex Systems*, 13(4), pp. 519-533. [doi:10.1142/S0219525910002694]

MAYO N. E., Brophy, J., Goldberg, M. S., Klein, M. B., Miller, S., Platt, R. W., and Ritchie, J. (2006) Peering at Peer Review Revealed High Degree of Chance Associated with Funding of Grant Application. *Journal of Clinical Epidemiology* 59, pp. 842-848. [doi:10.1016/j.jclinepi.2005.12.007]

MERTON, R. K. (1973) *The Sociology of Science. Theoretical and Empirical Investigations.* Chicago: University of Chicago Press.

NATURE (2006) Peer Review and Fraud. *Nature* 444, pp. 971-972. [doi:10.1038/444971b]

OBRECHT, M., Tibelius, K. and D'Aloiso, G. (2007) Examining the Value Added by Committee Discussion in the Review of Applications for Research Award. *Research Evaluation* 16 pp. 79-91. [doi:10.3152/095820207x223785]

PULVERER, B. (2010) Transparency Showcases Strength of Peer Review. *Nature* 468, pp. 29-31. [doi:10.1038/468029a]

SCHAFFER, A. (2009) America's Got Science Talent: The Biomedical Research Community Goes Bananas for $200 Million in Stimulus Funding." *Slate*, April 29. http://www.slate.com/articles/health_and_science/science/2009/04/americas_got_science_talent.html

SCHWARTZ, S. J. and Zamboanga, B. L. (2009) The Peer-Review and Editorial System. Ways to Fix Something that Might Be Broken. *Perspectives on Psychological Science*, 4(1), pp. 54-61. [doi:10.1111/j.1745-6924.2009.01106.x]

SMITH, R. (1997) Peer Review: Reform or Revolution? Time to Open the Black Box of Peer Review. *British Medicine Journal* 315, pp. 759-760. [doi:10.1136/bmj.315.7111.759]

SMITH, R. (2006) Peer Review. A Flawed Process at the Heart of Science and Journals. *Journal of the Royal Society of Medicine* 99, pp. 759-760. [doi:10.1258/jrsm.99.4.178]

SPIER, R. E. (2002) The History of Peer Review. *Trends in Biotechnology*, 20, 8, pp. 357-358. [doi:10.1016/S0167-7799(02)01985-6]

SQUAZZONI, F. (2010) Peering into Peer Review. *Sociologica*, 3, doi: 10.2383/33640, accessible at: http://www.sociologica.mulino.it/doi/10.2383/33640

SQUAZZONI, F. and Boero, R. (2010) 'Complexity-Friendly Policy Modelling'. In Ahrweiler, P. (ed.), Innovation in Complex Social Systems, Routledge, London, 2010, pp. 290-299

THURNER, S. and Hanel, R. (2010) Peer-Review in a World with Rational Scientists: Toward Selection of the Average. *arXiv:1008.4324v1*. http://arxiv.org/PS_cache/arxiv/pdf/1008/1008.4324v1.pdf

WARE, M. (2007) *Peer Review in Scholarly Journals: Perspective of the Scholarly Community - An International Study.* Bristol: Mark Ware Consulting. http://www.publishingresearch.net/documents/PeerReviewFullPRCReport-final.pdf

WEDEKIND, C. and Milinski, M. (2000) Cooperation Through Image Scoring in Humans. *Science,* 288, pp. 850-852. [doi:10.1126/science.288.5467.850]