OXFORD

Can transparency undermine peer review? A simulation model of scientist behavior under open peer review

Federico Bianchi[iD] **and Flaminio Squazzoni**[iD]*

Department of Social and Political Sciences, University of Milan, Via Conservatorio 7, Milan 20122, Italy
*Corresponding author. E-mail: flaminio.squazzoni@unimi.it

Abstract

Transparency and accountability are keywords in corporate business, politics, and science. As part of the open science movement, many journals have started to adopt forms of open peer review beyond the closed (single- or double-blind) standard model. However, there is contrasting evidence on the impact of these innovations on the quality of peer review. Furthermore, their long-term consequences on scientists' cooperation and competition are difficult to assess empirically. This paper aims to fill this gap by presenting an agent-based model that simulates competition and status dynamics between scholars in an artificial academic system. The results would suggest that if referees were sensitive to competition and status, the transparency achieved by open peer review could backfire on the quality of the process. Although only abstract and hypothetical, our findings suggest the importance of multidimensional values of peer review and the anonymity and confidentiality of the process.

Key words: peer review; science; reviewer bias; status; agent-based models.

1. Introduction

Peer review is an institutionalized process that encapsulates the self-organized, decentralized nature of the science system (Merton 1973). It is key to the social construction of scientific knowledge, with implications for the legitimacy and credibility of science for various stakeholders (Weller 2001; Lamont 2009a; Mallard et al. 2009; Strang and Siler 2015). While it is often said that peer review has a long and venerable tradition dating back centuries, the terms 'peer review' and 'referee' entered the vocabulary of science only about 70 years ago (e.g. Csiszar 2016; Batagelj et al. 2017; Grimaldo et al. 2018). For instance, leading journals, such as *Nature*, began using peer review only in 1973. The practice of journals involving external experts to judge the quality of manuscripts for publication reflected the expansion of scientific publications and the increasing specialization of scientific knowledge after World War II (Moxham and Fyfe 2017; Fyfe et al. 2020).

It is worth noting that whether single or double blind, one of the intrinsic characteristics of this practice has always been its confidential, closed form (Chubin and Hackett 1990; Rabesandratana 2013). For instance, in a recent study on peer review and editorial practices in sociology journals since the 1950s, Merriman (2020) has shown that regardless of adaptations and variations of certain practices, confidentiality and anonymity were deliberately established as a 'norm' to reflect an important value (i.e. 'social equality' between experts) and serve an intellectual aim rather than an organizational necessity. This was key to protect scientists against risks inherent in priority claims, reducing judgment bias, protecting experts' scrutiny from potential retaliation by manuscript authors, and distributing decision-making responsibility without compromising editorial autonomy (Biagioli 2002, 2012; Bornmann 2013; Pontille and Torny 2014; Moxham and Fyfe 2017; Fyfe et al. 2020).

Despite being instrumental in ensuring the quality of publications and credibility of scientific claims, peer review has always been contested, often also due to its lack of transparency (Alberts et al. 2008; Couzin-Frankel 2013; Csiszar 2016). For instance, after pioneering open access journals, Vitek Tracz, one of the most visionary and influential science publishers, chair of a conglomerate controlling more than 250 biology and medicine journals, concluded that: 'peer review is sick and collapsing under its own weight' and only defended by an 'irrational, almost religious belief'. He viewed the 'anonymity granted to reviewers' as the real problem, especially because this would give an unfair advantage to reviewers against authors in the race for priority of discoveries, e.g. deliberately delaying publication of competing groups (Rabesandratana 2013). Similar arguments have led certain journal editors to arrange the process so that revealing reviewers' identities could help remove these competitive advantages by neutralizing information asymmetries (Wolfram 2020; Ross-Hellauer 2022).

While certain dysfunctionalities of the conventional model have been questioned, such as editorial and reviewer bias (e.g. Alberts et al. 2008; Lee et al. 2013), two important exogenous factors seem to work against the gold standard of confidential peer review, i.e. technology progress and the changing public status of science. Note that these two factors have always influenced peer review (Spier 2002; Csiszar 2016;

Moxham and Fyfe 2017). When the first photocopiers became available in the late 1950s, editors began distributing copies of manuscripts by post to external experts. Similarly, later with the digitalization of documents and the advent of the Internet, the geographical extent of the referee pool dramatically expanded, communication improved, and time delay was reduced. Now, these technological advances have led to various experiments on open and post-peer review, including the online self-selection of reviewers, along with richer forms of commenting and engagement. According to many observers, this would make confidentiality irrelevant, if not often even impossible to guarantee (Tennant et al. 2017).

Let us now consider public status and all pressures applied to scholarly journals from their various stakeholders. While since World War II, peer review has responded to the need for credible and recognized standards of evaluation of science, socially justifying generous public funds, priorities are now transparency and the accountability of editorial practices. Under the impact of recent scandals, e.g. publications with deliberately fake data and manipulated conclusions in top scholarly journals (e.g. Couzin-Frankel 2013), anything looking 'confidential' related to editorial decisions is associated with monopolistic powers and organizational irresponsibility.

These factors have contributed to the development of experiments by scholarly journals, learned societies, and publishers on new forms of open peer review. For instance, *The EMBO Journal, eLife*, and *Frontiers* journals support pre-publication interaction and collaboration between reviewers, editors, and in some cases even authors, while *F1000 Research* has implemented advanced collaborative platforms to engage reviewers in post-publication open reviews. In short, these new forms can include (but are not limited to) publishing peer review reports, revealing reviewers' identities, post-reviewing manuscripts' self-organization, and establishing direct and public communication between reviewers and authors (Wicherts 2016; Ross-Hellauer 2017; Tennant et al. 2017; Wang and Tahamtan 2017).

However, there is no consensus among scholars on the positive/negative implications of these innovations (Ross-Hellauer 2022). While open peer review could stimulate fairer, constructive, and more positive reports, it seems that this does not increase the quality of reports, while also implying less willingness to participate by potential reviewers. For instance, a large-scale study on five Elsevier journals before and during their transition from closed to open peer review found no significant differences in referees' willingness to review, turn-around time, or recommendations. However, these results testified to the lack of confidence by reviewers when asked to reveal their identities. Indeed, only 8.1 per cent of 18,525 reports reported reviewer names, most probably when reviewers provided a positive recommendation of the manuscript (Bravo et al. 2019).

This would indicate that reviewers are sensitive to possible retaliation by authors and would preferably opt out when reporting on a weak manuscript. If we consider the current academic landscape in which competition for funds and academic career uncertainty is dramatic (Fang and Casadevall 2015; Edwards and Siddhartha 2017), it is possible that the desirable principles of transparency and accountability of open peer review could even clash with competing principles of fairness and disinterestedness, equally essential

in peer review. While examining the impact of open peer review on referees' willingness to review, turn-around, report quality, and recommendations is possible (e.g. Walsh et al. 2000; Bruce et al. 2016; Bravo et al. 2019), examining the effect of these innovations on aggregate, collective patterns requires time-dependent scales of observation, which are hardly achievable either observationally or experimentally.

To fill this gap, we built a theoretical simulation model that encapsulated certain abstract features of peer review and allowed us to develop hypothetical scenarios to examine non-trivial, counterfactual hypotheses on open peer review based on plausible assumptions on scientist behavior (Squazzoni and Takács 2011). As suggested by a recent review in this field (Feliciani et al. 2019), this approach is particularly valuable to support theoretical reflections on generative mechanisms of complex, evolving social patterns that quantitative research on aggregate variables cannot help to capture (Macy and Willer 2002; Epstein 2006). We used our simulation model to examine the level of structural coherence between our hypotheses on the evolution of the scientific context and its plausible effects on scientist behavior (e.g. Squazzoni 2012). Note that our model only aims to capture certain stylized characteristics of the scientific context and does not aim to determine a specific micro-macro causal mechanism of explanation (León-Medina 2017; Knight and Reed 2019). The design is to support a critical discussion on recent innovations in scholarly communication (Ramström 2018) and provide scenario analysis on possible evolution of the science system, which is hard to estimate empirically (Yun et al. 2015).

The paper is organized as follows: the following section presents our model, which draws on existing agent-based models of peer review (e.g. Squazzoni and Gandelli 2013; Bianchi et al. 2018). Here, we assumed that scientists could have different motivation when reviewing and are sensitive to status signals and competition. Indeed, besides being 'peers' of a cohesive community and directly and indirectly collaborating on knowledge development, scholars are also embedded in academic organizations with complex hierarchies of prestige, worth, and value, now explicitly quantified (e.g. H-index). While the competitive spirits of the 'homo academicus' have always been part of the scientific field (Merton 1973; Bourdieu 1990), the increasing importance of publication signals, the quantification of academic life, and the general conflation of 'value' and 'worth' into single indicators have institutionalized competition at within the current 'academic capitalism' (Fochler, 2016). Research has found that prestigious authors sometimes could benefit from certain advantages, such as having their manuscripts treated more quickly by editors (Sarigol et al. 2017), more favorable reviews when the identity of the author is revealed (Okike et al. 2016; Teplitskiy et al. 2018), or more favorable editorial decisions in case of contested reviews (Bravo et al. 2018). More generally, these distortions can be viewed as the effect of status dynamics among scholars, which encompass actual power differences as well as the social construction of prestige and reputation (Sorokin 1927; Blau 1964). While some of these distortions could be ideally reduced by opening internal journal information, others could also be exacerbated or have simply unpredictable consequences, such as retaliation or collusion among scholars in other occasions.

Any form or innovation introduced in peer review from scholarly journals cannot ignore the fundamental fact that scientists do not work in an institutional vacuum. Scientists are embedded in systems of standards, rewards, and norms that orient their perceptions of value and worth (Reitz 2017). We hypothesized that institutional contexts could affect scientist behaviors and examined these potential effects on very simple measures of quality and efficiency of peer review practices. The third section presents and discusses our results, while the last discusses certain implications of our findings on the current debate on peer review.

2. The model

Building on a previous theoretical model by Squazzoni and Gandelli (2013), we simulated an idealized scientific community of N scientists, who alternatively performed the roles of *authors* and *reviewers* of manuscripts for scholarly journals. When authors, scientists submit manuscripts for publication, while when reviewers, they evaluate a randomly assigned submitted manuscript. At each iteration of the model, half of the scientists are randomly assigned as authors, half as reviewers. For the sake of simplicity, we assumed that each author submits only one manuscript at each model iteration and that each manuscript is reviewed by one reviewer, who is required to assign a score to the manuscript. At the end of each iteration, a fixed proportion P of submissions is selected for publication based on the best scores assigned by reviewers.

We assumed that both authoring and reviewing are costly activities. Each scientist is provided with a variable amount of resources r_i, which they consume to perform their tasks. While a portion of r_i is fixed, each scientist is provided with a small, equal amount of resources (E) at the beginning of each iteration. At the end of each iteration, scientists are assigned extra-resources in case their submitted manuscripts are eventually selected for publication. This was to model a productivity-based resource allocation policy, mitigated by access to minimal resources to which all scientists are entitled independent of their performance, such as research infrastructure provided by their organizations, access to libraries, online repositories, and laboratories.

We assumed that each submitted manuscript has an intrinsic quality value. This depends on each author's *expected quality* ($\bar{q}_l$), which is calculated at the beginning of each iteration as a function of the author's resources, as follows:

$$\bar{q}_l = \frac{v \cdot r_i}{v \cdot r_i + 1} \tag{1}$$

where v indicates the velocity at which authors' expected quality varies according to the increase of author's resources. We set $v = 0.1$ (see Table 1) so that the expected quality of authors increased slowly according to their resources. This was to model scientists who had some inertia in adapting the quality of their manuscripts to their increased resources. Note that systematic parameter explorations showed that model outcomes were not significantly affected by the variability of v (see Supplementary Figures S1–S3). Then, the actual *submission quality* (q_m) of a manuscript is randomly drawn from a normal distribution as follows:

$$N(\bar{q}_l, \sigma = B \cdot \bar{q}_l) \tag{2}$$

Table 1. Initialization of model parameters.

Parameter	Description	Value
N	Number of scientists	240
$r_i\ (t = 0)$	Scientists' initial resources	0
E	Fixed resource gain	1
P	Proportion of accepted publications	0.25
B	Author bias factor	0.1
v	Velocity of best quality approximation	0.1
d	Discount factor on resources for unreliable reviews	0.5

where B is the *author bias factor*, defined as an exogenously fixed parameter that biases the link between author's resources (r_i) and submission quality (q_m). This was to model possible variability of quality across different submissions by the same author, due to idiosyncratic factors, e.g. variations in the originality of research ideas, delay in priority claims, etc. We set $B = 0.1$ (see Table 1) to introduce a small effect. Note that systematic parameter explorations showed that model outcomes were not significantly affected by the variability of B (see Supplementary Figures S4–S6). Moreover, we assumed that the *submission quality* also determines the amount of resources consumed to submit a manuscript: submitting better manuscripts requires authors to invest more resources.

When reviewing, scientists estimate the quality of their assigned manuscript. In order to model reviewing bias, the estimated quality ($s_{i,m}$) of a manuscript m by a reviewer i depends on the distance between a reviewer's *expected quality* as an author and the *submission quality* of the manuscript, as in the following equation:

$$S_{i,m} = \frac{r_i(1 + q_m - \bar{q}_i)}{2} \tag{3}$$

where r_i is the amount of resources of reviewer i, q_m is the submission quality of manuscript m, and $\bar{q}_i$ is the expected quality of reviewer i when submitting his/her own manuscripts.

Therefore, Equation (2) reflects the assumption that the cost of reviewing increases with the distance between a reviewer's skills and the skills ideally required to ensure the best possible review of the manuscript due to its quality. Similar to the use of author resources, we assumed that the amount of resources required for reviewing are also determined by the quality of the review, hence by $s_{i,m}$. Therefore, if reviewers are assigned a manuscript whose quality is close to that of their own submissions, we assumed that they must allocate 50 per cent of their resources to deliver their review. Here, we assumed that standards of quality and complexity of research are stratified and mutual understanding increases between experts with the same standards, which are reflected by the similarity of their level of resources. Furthermore, fewer resources are required if reviewers are assigned manuscripts of lower quality, while more resources are consumed if they are assigned manuscripts of higher quality. Moreover, reviewing expenses were inversely proportional to the reviewer's level of resources. This was to mirror the fact that top scientists usually require less time and effort for reviewing, as they are more experienced and skilled than average scientists. However, we linked effort and value as a proportion as when reviewing similar manuscripts, more productive scientists used more

resources than average scientists because their time was more costly.

In order to keep the model simple and focus on the scientists' behavior, we did not explicitly model a publishing landscape with different journals having different standards and selectivity (see Kovanis et al. 2016). We simply assumed a selective scholarly journal market. In our model, to mimic a draconian 'publish or perish' academic environment, authors of unpublished manuscripts lose all resources invested in the submission process. On the contrary, if published, author resources are multiplied by a value M, ranging between 1 (for most resourceful scientists) and 1.5 (for least resourceful scientists). This was to model different marginal benefits yielded from resource gains for scientists depending on their level of resources (ideally reflecting the number of their previous publications).

The assumption was that the marginal utility of each publication is higher at the beginning of a scientist's career than in later stages. For instance, the nth publication by already well-established scientists has less impact on their reputation and value than the first publications by less-established colleagues. More precisely, at the end of each iteration, all n published authors are ranked by level of resources according to an A_i sequence, where A_0 represents the published author with the lowest level of resources and A_{n-1} represents the published author with the highest level of resources. We calculated the multiplier M for scientist i as follows:

$$M(A_i) = g - \frac{i}{n-1}(g-1) = \frac{g(n-i-1)+i}{n-1} \qquad (4)$$

where g is the higher multiplier possible and i is the rank of an author by resources.

3. Model specification and simulation scenarios[1]

We created two simulation scenarios in which we assumed adaptive behavior by scientists and introduced different idealized peer review models: confidential peer review vs. open peer review. In the first one (*confidential peer review*), we assumed that authors and reviewers are anonymous during the process. We assumed that scientists could be more or less reliable as reviewers depending on their previous publication/rejection when authors (i.e. *indirect reciprocity*). For instance, if their manuscript was previously published, they could be more reliable when later casted as reviewers, whereas otherwise, they could retaliate by performing biased reviews. We assumed that in these cases, unreliable reviews required less resources from reviewers to be performed (i.e. the reviewing cost was decreased by a d factor) as if they were saving time and resources to concentrate on their manuscripts (Bianchi et al. 2018). We assessed the impact of this behavior against two baseline scenarios where scientists were either always (i.e. *fair*) or never reliable (i.e. *unreliable*). In the latter case, reviewers could either over- or under-rate the quality of manuscripts randomly.

In the second scenario (*open peer review*), we assumed that author and reviewer identities are revealed during the process. This means that scientists could use these signals to play dyadic 'TIT-for-TAT' moves: promoting previously positive or punishing previously negative reviewers when reviewing their manuscripts later (i.e. *direct reciprocity*). Note that we

assumed that reviewing was always costly in the *open peer review* model, because of the risk of retaliation related to identity disclosure and the fact that reviews could be published. Moreover, in the *status* scenario, we assumed that reviewers could be influenced by author status in that they are keen to pay deference to more established authors by being more positive while reviewing their submissions. In this scenario, whenever matched with a manuscript, reviewers in the lowest quartile of the resource distribution could estimate the status of their matched authors by comparing their respective resources (r_i). In case reviewers' resources were less than twice the amount of resources of manuscript authors, reviewers paid deference to authors' higher status by overrating their manuscript. We assessed the effect of this status-based mechanism against a scenario where 25 per cent of scientists were randomly selected to perform biased reviews (i.e. *bias*).

We simulated each scenario for 200 realizations. For each realization, we ran the model for 3,000 iterations in order to achieve robust outcome measurements by minimizing the effect of stochastic parameters and ensure full comparability of all scenarios. Note that time here does not reflect any real event, e.g. the number of manuscripts that scientists can realistically submit in their lifetime, but it is only a feature of the model.

Table 1 summarizes the model parameters and their initialization values. Table 2 shows the pseudo-code of the model. In order to test the sensitivity of model outcomes to varying initialization values, we conducted a systematic exploration of the model's parameter space, which showed no significant qualitative differences between scenarios compared to parameter values reported in Table 1 (see Supplementary material).

We calculated three outcome measurements at the end of each realization. We measured *publication bias* as the percentage of published manuscripts that should not have been published if reviewers had correctly estimated their actual *submission quality*. We then calculated *reviewing expenses* as the percentage of the total resources spent by reviewers over the total amount of resources spent by authors in each time step. Finally, we calculated the *average published quality* as the average *submission quality* of published articles.

5. Results

Figure 1 shows the impact of reviewer behavior on *publication bias* when peer review is confidential. Similar to Squazzoni and Gandelli (2013), the results showed that if scientists would follow reactive strategies while reviewing, i.e. reciprocating past experience as authors by being more or less fair when casted as reviewers, peer review cannot ensure that only the best submissions are published. Note that the outcome of this scenario is similar to a model of a purely random peer review. This would mean that confidentiality would ensure high quality of peer review only when scientists do not behave strategically when acting as reviewers. Compared with an optimal situation (*fair* scenario), in which all scientists behave fairly and try to promote only high-quality submissions, the percentage of bias with strategic reviewers in our simulations increased, on average, by almost 40 per cent. This would indicate that widespread concern about potential distortions of strategic referees could be well placed, especially considering

the current hyper-competitive landscape in which publication signals are so important (Edwards and Siddhartha 2017).

Table 2. Pseudo-algorithm of the model core. For specific details concerning different scenarios, see the model code at [ANONYMIZED].

Input: time t, number of iterations T, set of n agents, r,
Output: publication bias, average publication quality, top quality

1	initialize $t = 0$
2	initialize n agents
3	**for all** agents i:
4	$r_i \leftarrow 0$
5	# publications $\leftarrow 0$
6	**end for**
7	**while** $t < T$ **do**
8	**for all** agents i:
9	update r_i
10	compute $\overline{q_i}$
11	initialize $n/2$ reviewers
12	**for all** reviewers j:
13	initialize behavior (reliable vs. unreliable)
14	match with one random non-reviewer agent i
15	**end for**
16	**for all** non-reviewers i:
17	compute q_m
18	compute spent resources
19	**end for**
20	**for all** reviewers j:
21	compute $s_{i,m}$
22	**end for**
23	**end for**
24	rank submitted manuscripts m by $s_{i,m}$
25	**for all** non-reviewers i:
26	**if** m ranked among top $Pn/2$ manuscripts **do**
27	# publications $\leftarrow$ # publications $+ 1$
28	**end if**
29	**end for**
30	**end while**

Figure 2 shows a comparison between *confidential* and *open peer review* when scientists behave strategically (i.e. *indirect reciprocity* vs. *direct reciprocity*). Here, it is important to note that while confidentiality implied only indirect reciprocity motives (i.e. whenever rejected, scientists retaliated against peer review by providing unfair reviews, saving time and resources to prepare their own manuscripts), *open peer review* can induce scientists to play direct reciprocity strategies. Indeed, reviewers could recognize authors who previously reviewed (either positively or negatively) their submissions and so play a direct 'TIT-for-TAT' cooperation strategy. It is important to note here that previous surveys on scientists involved in open peer review considered this reactive behavior as one of the most critical inconveniences of open peer review (e.g. Ross-Hellauer et al. 2017). The results showed that open peer review would cause a dramatic increase in the publication bias compared to confidential, reaching a level of 70 per cent.

Figure 3 compares various *open peer review* scenarios. While in the previous scenario, we considered reciprocity strategies that were enacted simply by competitive signals, here we considered more sophisticated signals. We assumed that transparency led reviewers to estimate the status of the manuscript authors they were matched with while knowing at the same time that authors could do the same in their turn at the end of the process. This is especially relevant when considering all potential author–reviewer matching, including low-status reviewers, e.g. Ph.D. students or post-docs, being asked to review submissions by authors of higher status. As previously mentioned, as a benchmark, we created a scenario in which status perception was not linearly linked to scientist productivity in that a variety of possible distortions exist beyond productivity-based status, e.g. academic seniority, power, or personal relationships.
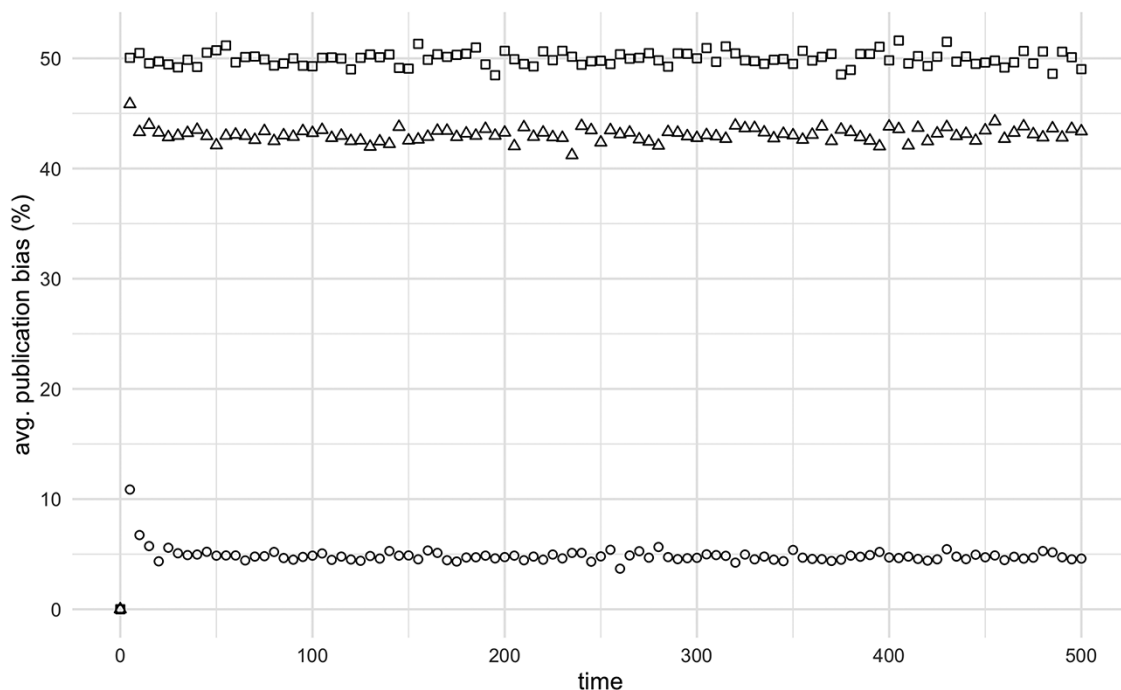
Figure 1. Impact of reviewer behavior on *publication bias* in *confidential peer review* over time. Circles: *fair*; squares: *unreliable*; triangles: *indirect reciprocity*. (Values averaged over 200 realizations.)
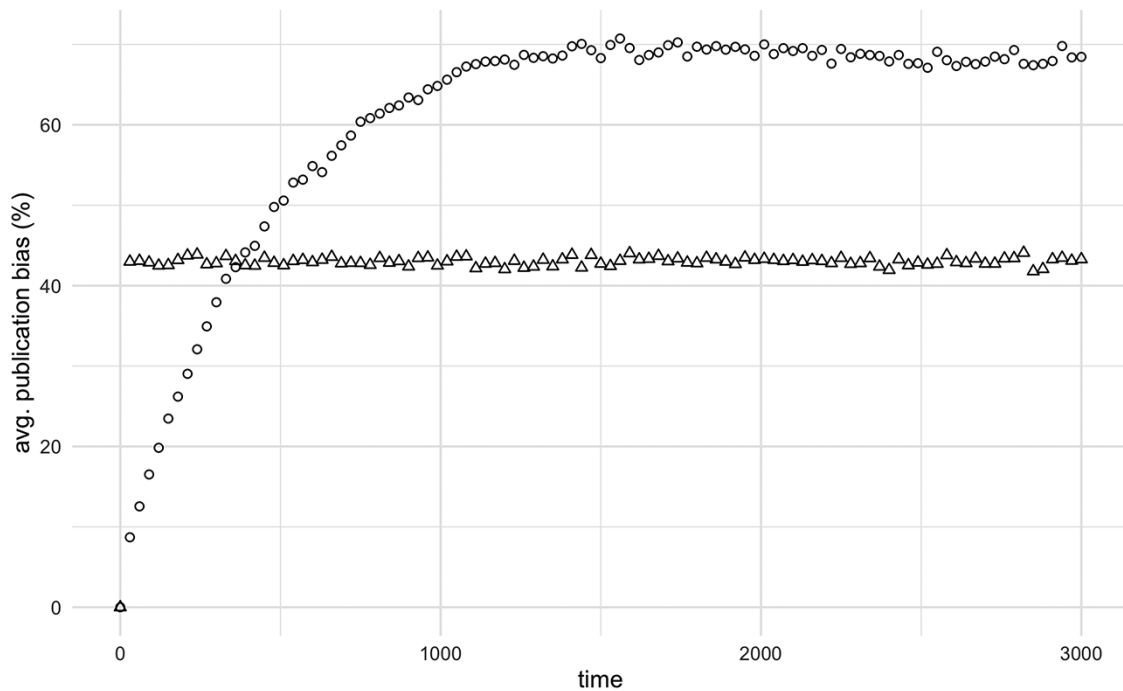
Figure 2. Impact of scientist reciprocity strategies on *publication bias* over time in *confidential* vs. *open peer review*. Triangles: indirect reciprocity (*confidential peer review*); circles: direct reciprocity (*open peer review*). (Values averaged over 200 realizations.)
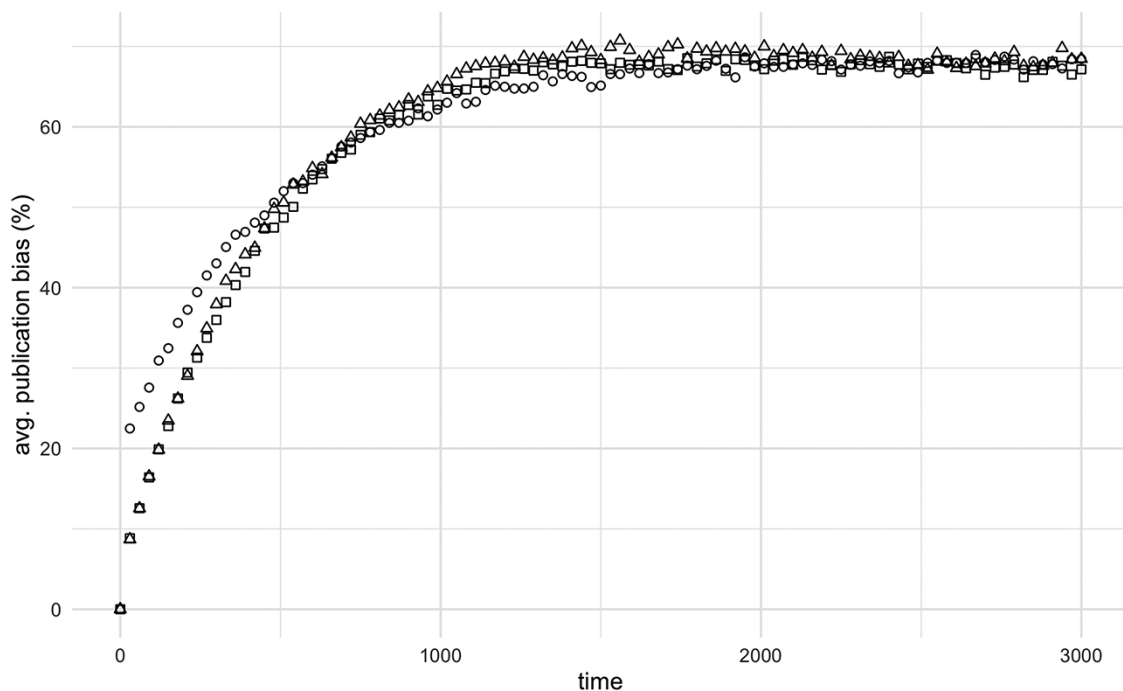

Figure 3. Impact of author status on *publication bias* over time in confidential vs. *open peer review*. Triangles: *direct reciprocity*; squares: *status*; circles: *bias*. (Values averaged over 100 realizations.)

The results showed that regardless of the source of status, *publication bias* tended to converge to the level of 70 per cent mentioned above. This had a detrimental effect on the average quality of publications in all scenarios (see Fig. 4).

It is worth noting here that we assumed a relatively fair scenario in that status affects only scientists' behavior during peer review. We did not contemplate scenarios in which reviewers

could be retaliated to by the punished authors more academically, e.g. high-status scientists refusing to hire colleagues who provided a negative review on a previous manuscript of theirs.

Figure 5 shows a comparison of simulation scenarios on reviewing expenses, which was a measure of system efficiency. The results showed that *open peer review* was more demanding in terms of resources at a system level. Indeed, this caused
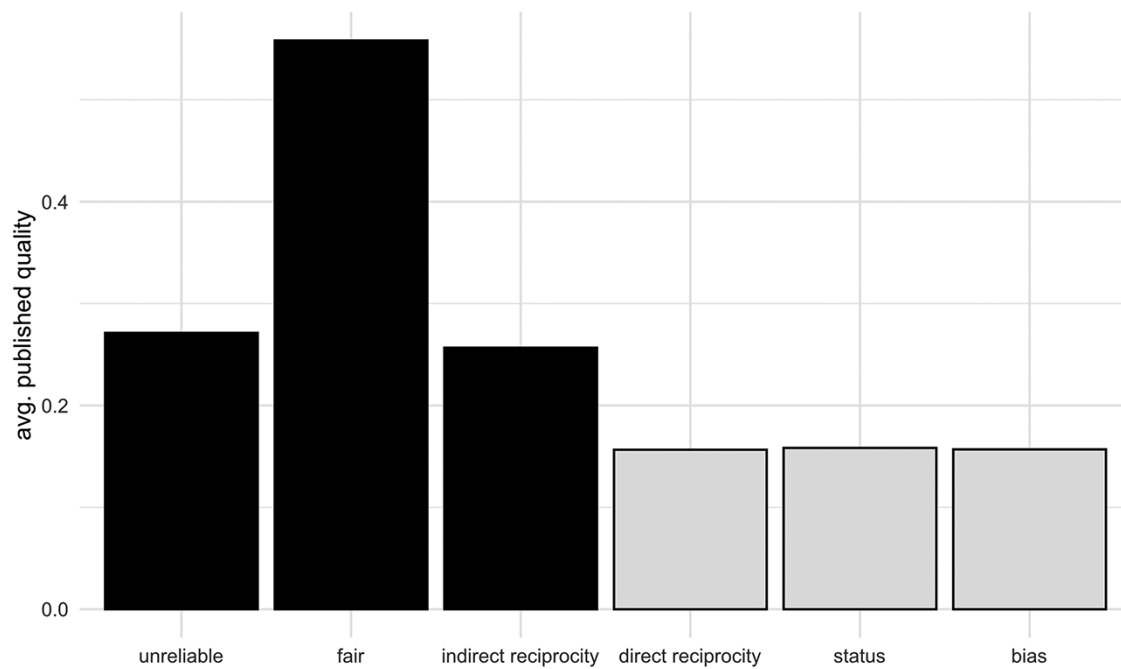
Figure 4. Impact of reviewer behavior on the average quality of published papers under different peer review models. Black: confidential peer review; white: open peer review. (Values calculated at last iteration and averaged over 200 realizations.)
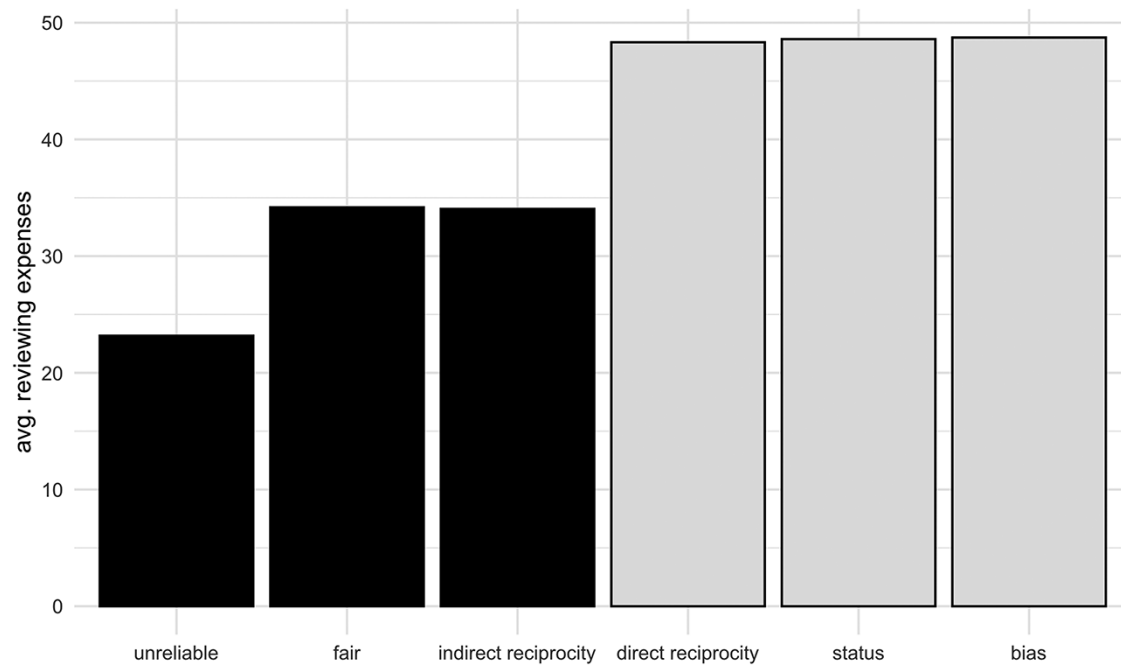

Figure 5. Impact of reviewer behavior on reviewing expenses under different peer review models. Black: confidential peer review; white: open peer review. (Values calculated at last iteration and averaged over 200 realizations.)

a resource drain from research almost twice the amount of resources consumed in random peer review (*unreliable*) and about 10 per cent more of the other scenarios in *confidential peer review*. This would confirm previous experimental findings on more time and resources required by reviewers when under open peer review (e.g. Bruce et al. 2016). Note that the low resource waste of confidential peer review when reviewers had a constant probability of being unreliable depends on the assumption that being unreliable means providing a random

opinion on a manuscript without spending time in reading it carefully (e.g. Squazzoni and Gandelli 2013; Bianchi et al. 2018).

Finally, when looking at resource distribution at the system level and comparing all scenarios, we found that whenever reviewers are sensitive to reciprocity motivations, open peer review would decrease the inequality in resource distribution between scientists by 10 per cent compared to confidential peer review (Gini index of 0.25 vs. 0.35; see Supplementary

Figure S4). Coherently with previous simulation findings (Squazzoni and Gandelli 2013), and given the draconian 'publish or perish' academic environment assumed in the model where all resources invested in manuscripts are lost whenever they are rejected, inequality would increase only when reviewers are fair and under confidential peer review. In short, in our hypothetical and artificial scenarios, open peer review would increase publication bias but decrease inequality.

6. Discussion and conclusions

Although only abstract and highly hypothetical, our simulation experiments would suggest that open peer review could even possibly undermine the quality and efficiency of the process. This would be the case if scientists, when reviewing, adapt their behavior to personal success or failure and are sensitive to status. While the scientific community is made up of 'peers', i.e. ideally equal, disinterested members, the social and organizational process characterizing scientific competition for priority and rewards entails the emergence of implicit hierarchical structures in which status differences are important signals of value and worth (Cole and Cole 1973; Bourdieu 1990; Burris 2004; Reitz 2017). The fact that these structures could in principle determine disciplinary balkanization, power, and conflicts does not devalue their important role in fostering cumulative knowledge by providing coordination signals to scientists (what is relevant, what is not) and epistemic standards of worth and value to science stakeholders (what is scientific knowledge, what is not). Unfortunately, open science supporters tend to neglect the importance of these contextual factors while concentrating on the ultimate, superior goals of transparency and accountability (Ross-Hellauer 2022). However, in complex institutional environments, the co-existence of multiple values typically creates tensions and conflicts, which need to be considered, publicly discussed, and collectively regulated (Minssen et al. 2020), not to mention the possible co-existence of different normative equilibria in various communities (Martin 2016).

It is reasonable to expect that the effect of these complex social processes could influence the current practices of peer review in scholarly journals, especially if the 'veil of anonymity', typically protecting these practices, is removed (Bravo et al. 2018). It is worth noting that whenever such a veil is removed, not only do authors and reviewers know their respective counterparts, but authors also know that reviewers understand this while reviewing and reviewers could use reviewing to send implicitly collusive or conflictual signals. At the same time, reviewers lose the possibility of claiming their ignorance about authors' identities as an excuse, if retaliated to by susceptible authors (e.g. Mäs and Opp 2016).

The practice of double-blind peer review in social science journals is in the first place a means to protect reviewers from possible retaliation, thus ensuring the independence of judgment (Pontille and Torny 2014; Horbach and Halffman 2018; Merriman 2020). The fact that authors are now easily traceable on Google or other Internet-based sources—which has made many analysts suggest abandoning double-blind peer review as a mere ritual—is completely irrelevant when dismissing this practice (Weicher 2008; Nobarany & Booth 2017). It is worth noting that transparency can be especially detrimental in cases of young scholars who could be sensitive to the risk of retaliation when asked to review work by more

advanced senior scholars (e.g. Wang et al. 2016). As argued by Flaherty (2016) while reflecting on his experience as editor of many sociology journals: 'How many of us will truthfully point out the flaws in a colleague's manuscript when he or she will know our identity? Will a junior scholar reject the manuscript submitted by a senior scholar?' In our view, these complex status effects could also explain the low level of consensus that open peer review has among the humanities and social sciences, which are more sensitive to inequality, status, and justice (Ross-Hellauer et al. 2017).

Despite being highly abstract and hypothetical, our findings suggest that science is a complex evaluation system of practice characterized by a multidimensional space of 'values' and worth, with potential conflicts and trade-offs (Mallard, Lamont & Guetzkow 2009; Lamont 2012; Martin 2016). For instance, while improving transparency is surely a valuable goal, genuinely inspiring many laudable experiments on peer review, this could come at the cost of reducing the independence of judgment, criticism, and the search for truth, always central to the most important pillars of science as an institutional system (Merton 1973).

Although abstract and highly hypothetical, our simulation helps to stimulate the call to reconsider these aspects while testing manipulations of peer review. Furthermore, while measuring the effects of these manipulations is in principle possible with in-journal data (e.g. Bravo et al. 2019), tracing long-term consequences of scientists' reactions would require cross-journal data and consistent identity tracing procedures, as scientists review manuscripts for different journals: in short, a very arduous endeavor. This is where simulation models can help by providing scenario analysis that exploits available knowledge to explore probabilistic future trends with all due caveats concerning the weak realism and attention to contexts (Gilbert et al. 2018; Feliciani et al. 2019).

Following previous research, at first glance our simulations would lead to a dramatic conclusion, i.e. if scientists behave adaptively in peer review, a random system could be functional to reduce the cost of reviewing (e.g. Squazzoni and Gandelli 2013). This is true, according to our model, which has been intentionally built to consider extreme (hypothetical) scenarios. However, the problem is that it is hard to believe that even ideally, scientists would behave purely randomly. Secondly, the fact that scientists follow adaptive behavior makes them sensitive to incentives and regulations, which are the real levers that scientific associations, publishers, and journals have at their disposal to perform experiments that can improve the current situation. This is relevant even when contemplating reforms that are more coherent to the complex, multidimensional space of values and context-specific set of practices characterizing the academic community. Obviously, while artificial experiments can stimulate debate by analogy, their findings cannot inform policy or interventions and would require empirical test to support any generalization (Edmonds et al. 2019).

Finally, our simulations are only theoretical and speculative as more evidence-based, empirically informed research would be necessary to reflect on the various causes of these strategic behaviors, including social influence processes, learning, social norms, and more complex institutional dynamics (Feliciani et al. 2019). While the advantage of agent-based models is that they permit us to look at stylized cases and draw theoretical reflections from premises in logically consistent

ways, integrating these models with empirical data is the key both to calibrate agent behavior and incorporate contextual factors. Here, a set of well-design case studies comparing similar journals with different peer review models and tracing author/referee previous connections would help to calibrate sensitive model parameters, reduce abstract theoretical specifications, and provide a more context-specific picture of the behavioral dynamics underlying peer review outcomes. Unfortunately, there are still obstacles against large-scale data sharing from journals to study peer review due to confidentiality, lack of incentives to publishers, and lack of collaborative infrastructure (Squazzoni et al. 2020), which all need to be confronted.

Supplementary data

Supplementary data is available at *SCIPOL Journal* online.

Notes

1. The model has been built using NetLogo 6.2. The model code is available here: https://www.comses.net/codebases/3d99eb9f-ae4f-42d0-8c58-9d28757161c0/releases/1.0.0/.

References

Alberts, B., Hanson, B., and Kelner, K. L. (2008) 'Reviewing Peer Review', *Science*, 321: 15.

Batagelj, V., Ferligoj, A., and Squazzoni, F. (2017) 'The Emergence of A Field: A Network Analysis of Research on Peer Review', *Scientometrics*, 113: 503–32.

Biagioli, M. (2002) 'From Book Censorship to Academic Peer Review', *Emergences: Journal for the Study of Media & Composite Cultures*, 12: 11–45.

——— (2012) 'From Ciphers to Confidentiality: Secrecy, Openness and Priority in Science', *The British Journal for the History of Science*, 45: 213–33.

Bianchi, F., Grimaldo, F., Bravo, G., et al. (2018) 'The Peer Review Game: An Agent-based Model of Scientists Facing Resource Constraints and Institutional Pressures', *Scientometrics*, 116: 1401–20.

Blau, P. M. (1964) *Exchange and Power in Social Life*. New Brunswick: Transactions.

Bornmann, L. (2013) 'Scientific Peer Review', *Information Science & Technology*, 45: 197–245.

Bourdieu, P. (1990) *Homo Academicus*. Stanford: Stanford University Press.

Bravo, G., Farjam, M., Grimaldo, F., et al. (2018) 'Hidden Connections: Network Effects on Editorial Decisions in Four Computer Science Journals', *Journal of Informetrics*, 21: 101–12.

Bravo, G., Grimaldo, F., López-Iñesta, E., et al. (2019) 'The Effect of Publishing Peer Review Reports on Referee Behavior in Five Scholarly Journals', *Nature Communications*, 10: 322.

Bruce, R., Chauvin, A., Trinquart, L., et al. (2016) 'Impact of Interventions to Improve the Quality of Peer Review of Biomedical Journals: A Systematic Review and Meta-analysis', *BMC Medicine*, 14: 85.

Burris, V. (2004) 'The Academic Caste System: Prestige Hierarchies in PhD Exchange Networks', *American Sociological Review*, 69: 239–64.

Chubin, D. E. and Hackett, E. J. (1990) *Peerless Science*. Albany: State University of New York Press.

Cole, J. R. and Cole, S. (1973) *Social Stratification in Science*. Chicago: University of Chicago Press.

Couzin-Frankel, J. (2013) 'Secretive and Subjective, Peer Review Proves Resistant to Study', *Science*, 342: 1331.

Csiszar, A. (2016) 'Troubled from the Start', *Nature*, 532: 306–8.

Edmonds, B., Le Page, C., Bithell, M., et al. (2019) 'Different Modelling Purposes', *Journal of Artificial Societies and Social Simulation*, 22: 6.

Edwards, M. A. and Siddhartha, R. (2017) 'Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition', *Environmental Science and Engineering*, 34: 51–61.

Epstein, J. (2006) *Generative Social Science. Studies in Agent-Based Modeling*. Princeton: Princeton University Press.

Fang, F. C. and Casadevall, A. (2015) 'A Competitive Science: Is Competition Ruining Science?' *Infective Immunology*, 83: 1229–33.

Feliciani, T., Luo, J., Ma, L., et al. (2019) 'A Scoping Review of Simulation Models of Peer Review', *Scientometrics*, 121: 555–94.

Flaherty, M. G. (2016) 'Sociology as a Conversation: The Present Circumstances and Future Prospects of Peer Review', *The American Sociologist*, 47: 253–63.

Fochler, M. (2016) 'Variants of Epistemic Capitalism: Knowledge Production and the Accumulation of Worth in Commercial Biotechnology and the Academic Life Sciences', *Science, Technology & Human Values*, 41: 922–48.

Fyfe, A., Squazzoni, F., Torny, D., et al. (2020) 'Managing the Growth of Peer Review at the Royal Society Journals, 1865-1965', *Science, Technology & Human Values*, 45: 405–29.

Gilbert, N., Ahrweiler, P., Barbrook-Johnson, P., et al. (2018) 'Computational Modelling of Public Policy: Reflections on Practice', *Journal of Artificial Societies and Social Simulation*, 21: 14.

Grimaldo, F., Marusic, A., and Squazzoni, F. (2018) 'Fragments of Peer Review: A Quantitative Analysis of the Literature (1969-2015)', *PLOS ONE*, 13: 1–14.

Horbach, S. and Halffman, W. (2018) 'The Changing Forms and Expectations of Peer Review', *Research Integrity & Peer Review*, 3: 8.

Knight, C. R. and Reed, I. A. (2019) 'Meaning and Modularity: The Multivalence of "Mechanism" in Sociological Explanation', *Sociological Theory*, 37: 234–56.

Kovanis, M., Porcher, R., Ravaud, P., et al. (2016) 'The global burden of journal peer review in the biomedical literature: Strong imbalance in the collective enterprise', *PLOS One*, 11: e0166387.

Lamont, M. (2009a) *How Professors Think. Inside the Curious World of Academic Judgement*. Cambridge: Harvard University Press.

——— (2009b) *How Professor Think. Inside the Curious World of Academic Judgement*. Cambridge: Harvard University Press.

——— (2012) 'Toward a Comparative Sociology of Valuation and Evaluation', *Annual Review of Sociology*, 38: 201–11.

Lee, C. J., Sugimoto, C. R., Zhang, G., et al. (2013) 'Bias in Peer Review', *Journal of the American Society for Information Science and Technology*, 64: 2–17.

León-Medina, F. (2017) 'Analytical Sociology and Agent-based Modeling: Is *Generative Sufficiency* Sufficient?' *Sociological Theory*, 35: 157–78.

Macy, M. and Willer, R. (2002) 'From Factors to Actors: Computational Sociology and Agent-based Modeling', *Annual Review of Sociology*, 28: 143–66.

Mallard, G., Lamont, M., and Guetzkow, J. (2009) 'Fairness as Appropriateness: Negotiating Epistemological Differences in Peer Review', *Science, Technology & Human Values*, 34: 573–606.

Martin, B. R. (2016) 'Twenty Challenges for Innovation Studies', *Science & Public Policy*, 43: 432–50.

Mäs, M. and Opp, K.-D. (2016) 'When Is Ignorance Bliss? Disclosing True Information and Cascades of Norm Violation in Networks', *Social Networks*, 47: 116–29.

Merriman, B. (2020) 'Peer Review as an Evolving Response to Organizational Constraint: Evidence from Sociology Journals, 1952–2018', *The American Sociologist*, 52: 341–66.

Merton, R. K. (1973) *The Sociology of Science*. Chicago: University of Chicago Press.

Minssen, T., Rajam, N., and Bogers, M. (2020) 'Clinical Trial Data Transparency and GDPR Compliance: Implications for Data Sharing and Open Innovation', *Science & Public Policy*, 47: 616–26.

Moxham, N. J. and Fyfe, A. (2017) 'The Royal Society and the Prehistory of Peer Review, 1665-1965', *The Historical Journal*, 61: 863–89.

Nobarany, S. and Booth, K. S. (2017) 'Understanding and Supporting Anonymity Policies in Peer Review', *Journal of the Association for Information Science and Technology*, 68: 957–71.

Okike, K., Hug, K. T., Kocher, M. S., et al. (2016) 'Single-blind vs Double-blind Peer Review in the Setting of Author Prestige', *JAMA*, 316: 1315–6.

Pontille, D. and Torny, D. (2014) 'The Blind Shall See! The Question of Anonomity in Journal Peer Review', *ADA: A Journal of Gender, New Media, and Technology*, 4.

Rabesandratana, T. (2013) 'The Seer of Science Publishing', *Science*, 342: 66–7.

Ramström, G. (2018) 'Coleman's Boat Revisited: Causal Sequences and the Micro-macro Link', *Sociological Theory*, 36: 368–91.

Reitz, T. (2017) 'Academic Hierarchies in Neo-feudal Capitalism: How Status Competition Processes Trust and Facilitates the Appropriation of Knowledge', *Higher Education*, 73: 871–86.

Ross-Hellauer, T. (2017) 'What Is Open Peer Review? A Systematic Review', *F1000 Research*, 6: 588.

——— (2022) 'Open Science, Done Wrong, Will Compound Inequities', *Nature*, 603: 363.

Ross-Hellauer, T., Deppe, A., and Schmidt, B. (2017) 'Survey on Open Peer Review. Attitudes and Experience Amongst Editors, Authors and Reviewers', *PLOS ONE*, 12: 1–28.

Sarigol, E., Garcia, D., Scholtes, I., et al. (2017) 'Quantifying the Effect of Editor-Author Relations on Manuscript Handling Times', *Scientometrics*, 113: 609–31.

Sorokin, P. A. (1927) *Social Mobility*. New York: Harper.

Spier, R. (2002) 'The History of the Peer Review Process', *Trends in Biotechnology*, 20: 357–8.

Squazzoni, F. (2012) *Agent-Based Computational Sociology*. Hoboken: Wiley & Sons.

Squazzoni, F., Arhweiler, P., et al. (2020) 'Unlock Way to Share Data on Peer Review', *Nature*, 578: 512–4.

Squazzoni, F. and Gandelli, C. (2013) 'Opening the Black Box of Peer Review. An Agent-based Model of Scientist Behaviour', *Journal of Artificial Societies and Social Simulation*, 16: 3.

Squazzoni, F. and Takács, K. (2011) 'Social Simulation That "Peers into Peer Review"', *Journal of Artificial Societies and Social Simulation*, 14: 3.

Strang, D. and Siler, K. (2015) 'Revising as Reframing: Original Submissions versus Published Papers in Administrative Science Quarterly, 2005 to 2009', *Sociological Theory*, 33: 71–96.

Tennant, J., et al. (2017) 'A Multi-disciplinary Perspective on Emergent and Future Innovations in Peer Review', *F1000 Research*, 6: 1151.

Teplitskiy, M., Acuna, D., Elamrani-Raoult, A., et al. (2018) 'The Sociology of Scientific Validity: How Professional Networks Shape Judgement in Peer Review', *Research Policy*, 47: 1825–41.

Walsh, E., Rooney, M., Appleby, L., et al. (2000) 'Open Peer Review: A Randomized Controlled Trial', *British Journal of Psychiatry*, 176: 47–51.

Wang, P., Hoyt, J., Pölsch, U., et al. (2016) 'The Last Frontier in Open Science: Will Open Peer Review Transform Scientific and Scholarly Publishing?' *Proceedings of the Association for Information Science and Technology*, 53: 1–4.

Wang, P. and Tahamtan, I. (2017) 'The State-of-the-Art of Open Peer Review: Early Adopters', *Proceedings of the American Society for Information Science and Technology*, 54: 819–20.

Weicher, M. (2008) 'Peer Review and Secrecy in the "Information Age"', *Proceedings of the American Society for Information Science and Technology*, 45: 1–12.

Weller, A. C. (2001) *Editorial Peer Review: Its Strengths and Weaknesses*. Medford: American Society for Information Science and Technology.

Wicherts, J. M. (2016) 'Peer Review Quality and Transparency of the Peer Review Process in Open Access and Subscription Journals', *PLOS ONE*, 11: 1–19.

Wolfram, D., Wang, P., Hembree, A., et al. (2020) 'Open Peer Review: Promoting Transparency in Open Science', *Scientometrics*, 125: 1033–51.

Yun, J. J. H., Won, D. K., Hwang, B., et al. (2015) 'Analysing and Simulating the Effects of Open Innovation Policies: Application of the Results to Cambodia', *Science & Public Policy*, 42: 743–60.